AUTOMATION IN RAPID TRANSIT CONTROL: The State of the Art

by

G. A. Abrahamsohn
Senior Research Officer

with

G. Heti
Project Research Officer

H. K. Chan
Research Officer

Systems Research and Development Branch

September 1977

Ontario

Ministry of
Transportation and
Communications

Research and
Development
Division

ABSTRACT

The purpose of this report is twofold. First and foremost, it is intended to serve as a "primer" on automatic train control for rapid transit systems. Secondly, the report is to provide an insight into the potential and realizable benefits of automatic train control. Constraining factors and considerations are discussed. A special attempt is made to introduce practical considerations into the assessment of automatic train control for rapid transit systems. It is hoped that the information and the considerations brought forward in this report will aid in the assessment of future systems.

CONTENTS

GLOSSARY OF TERMS

Aspect:	The visual indication presented to an approaching train by a wayside signal.
Audio frequency track circuit:	A track circuit energized by an electrical current alternating in the audio frequency range (20 to 20,000 Hz).
Automatic Train Control:	The use of automation to conduct all or most of the train control functions. The acronyms ATP (automatic train protection), ATO (automatic train operation), and ATS (automatic train supervision) denote groups of automated functions.
Block:	A length of track of defined limits, the use of which is governed by block signals, cab signals, or both.
Cab Signal System:	A signal system whereby block condition and the prevailing civil speed commands are transmitted and displayed directly within the train cab.
Closed Loop Principle:	The principle of control system design in which the response of a system is continuously compared to the controlling signal to generate an error signal.
Coast:	The moving of a vehicle or train where the propulsion is inactive.
Consist:	The number and type of cars making up a train.
Dwell:	The elapsed time from the time the train stops moving in a station until the instant it resumes moving.
Fail-Safe:	A characteristic of a system which ensures that a fault of any element affecting safety will cause the system to revert to a state that is known to be safe.
Frequency Shift-Keyed (FSK):	Modulation of the track signal between two or more discrete frequencies to convey information.

Headway:	The time separation between two trains.
Insulated Joint:	A joint placed between abutting rail ends to insulate them from each other electrically.
Interlocking:	Arrangement of signals and signal appliances so interconnected that their movement must succeed each other in proper sequence.
Jerk:	The rate of change of acceleration.
Married Pair:	Two semipermanently coupled cars that share certain essential components and are usually operated as a unit.
Mean Time Between Failures (MTBF):	Average time that a system or component will operate without failure or malfunction.
Mean Time to Restore (MTTR):	The average time required to restore a system or component to operation after a failure.
Overspeed Control:	Portion of carborne ATC system responsible for enforcing speed limits safely.
Peak Period:	The period during a weekday when system demand is highest. Usually 7:30 - 9:30 a.m. and 4:30 - 6:30 p.m.
Shunt:	A conductor joining two points in an electrical circuit so as to form a parallel or alternate current path.
Shunting Sensitivity:	The maximum impedance that, when placed at the most adverse shunting location, will cause the track circuit to indicate the presence of a train.
Signal:	Means of communicating direction or warning.

Speed Regulator: An onboard subsystem usually part of the ATO that controls acceleration and braking to cause the train to reach and maintain the desired speed.

Track Circuit: A detection device which provides information about the location of the trains at all times.

Train: A consist of one or more vehicles combined into an operating unit.

Trip Stop: A mechanical arm, located on the wayside, that can initiate an emergency brake on a train that passes it.

1. INTRODUCTION

The purpose of this report is twofold. First and foremost, it is intended to serve as a "primer" on automatic train control for rapid transit systems. Secondly, the report is to provide an insight into the potential and realizable benefits of automatic train control. Constraining factors and considerations are discussed.

Chapter 2 briefly outlines the main features of rapid transit systems as far as they apply to this report. Minimum Line-, Station-, and Turn-Around-headways are derived. A point of reference is established and some practical system considerations and operational concepts introduced. Chapter 3 summarizes the functional requirements for automatic train control systems. The usual division into automatic train protection (ATP), automatic train operation (ATO), and automatic train supervision (ATS) is made and the functional requirements for each established. Chapter 4 introduces the concept of signalling in rapid transit systems. An attempt is made to define the concepts in simple terms and to show the basis for the various signalling levels and the evolution to automatic train control.

Chapter 5 defines the concept of block systems and discusses hardware considerations and applications. It also deals with headway calculations for various signalling systems and train parameters.

In Chapter 6, the features of various train control systems are described. For automatic train operation systems (ATO) the division between narrow-band and broad-band systems is made. Advantages and disadvantages of both system architectures are described.

Chapter 7 offers a simplified cost/benefit analysis of automatic train operation in rapid transit systems. The potential benefits are balanced against system constraints which hinder the full realization of the benefits. An attempt is

made to quantify the benefits accurately but the system specificity of the benefits renders this undertaking illustrative rather than absolute.

Chapter 8 contains the summary and conclusions.

Parts of this report deal with technical factors and parts with economic considerations and readers with specialized interests may wish to skip those chapters that are not of direct interest to them. Those readers interested in the cost/benefit aspects of automatic train control should read Chapters 2 and 7. Those interested in gaining a basic understanding of present day train control technology should find Chapters 3, 4, 5 and 6 of interest.

A special attempt is made to introduce practical considerations into the assessment of automatic train control for rapid transit systems. Inadequate attention to the factors which limit the full realization of the expected benefits from automation leads to disappointing operational experience. Only a detailed consideration of all aspects of automatic train control and a system design which helps, rather than hinders, the realization of automatic train control benefits will result in a cost-effective automatic train control system. It must be kept in mind that the cost-effectiveness analysis represents a "snapshot" in the continuum of time. Conclusions which are true today may be overtaken by events of tomorrow. It is, therefore, important that each system be evaluated on its own merit. It is hoped that the information and the considerations brought forward in this report will aid in the assessment of future systems.

2. RAPID TRANSIT SYSTEMS

2.1 Introduction

The intent of this Chapter is to set the stage for the report. Rapid transit systems are characterized as steel-wheels-on-steel-rail systems operating on exclusive rights-of-way with on-line stations. Generally, such systems will have maximum speeds of 10 m/s to 30 m/s, capacities over 15 000 passengers per hour per direction and average station spacings between 500 m and 3000 m.

Interaction between various system parameters such as headway, train length, speed, and capacity are examined. Constraints on these parameters are given. Headway limitations at various points in the system are pointed out.

Finally, the possible role of automation is briefly examined and considerations of cost, performance, and safety are introduced.

2.2 System Constraints

To make the results of this report as general as possible, only three major system constraints are imposed. The first constraint restricts the study to vehicles with steel wheels on steel rails; the second limits the study to systems with fully reserved rights-of-way; and the third limitation is the consideration of systems with on-line stations only.

2.2.1 Vehicle on Rails

The need to restrict the study to steel-wheels-on-steel-rail systems is necessitated by the great number of variables that non rail guidance and suspensions introduce. For example, values of maximum acceleration and service braking are limited to 2.0 m/s^2 in the report. These are typical of rail systems but may not be appropriate for non rail systems. Switching

and branching techniques, operating strategies, and station designs are also different for rail-bound and non rail-bound systems.

2.2.2. On-Line Stations

With on-line stations, all vehicles travelling in one direction are forced to travel the same track which passes through the station. Vehicles may not pass one another. The flow of trains is restricted to the rate at which trains can pass through the station.

2.2.3. Exclusive Right-of-Way

In order to achieve reasonably high average speeds, a transit system needs separation from competing or interfering traffic. Schedules can only be maintained if unexpected delays are eliminated.

Level crossings are often unavoidable. Such intrusions on the system right-of-way may be acceptable if designed in such a way as to give preference to rail traffic, but this unavoidably results in delays in cross traffic and such delays become unacceptable where the rail system operates at headways under approximately three minutes.

Transit systems which operate at headways under three minutes, therefore, require fully exclusive rights-of-way with no cross traffic.

The analysis which follows assumes that an exclusive right-of-way system is being considered.

2.3 System Parameter Interaction

The parameter of fundamental interest to transit planners and designers is the capacity that a system can provide. Capacity may be defined as the number of passenger places passing a point on the track, in one direction, per unit time.

In equation form:

$$C = \frac{3600\ N_v K_t L_v}{H} \tag{2.1}$$

where C = line capacity (passengers/hour)

L_v = vehicle length (metres)

H = headway (seconds)

K_t = vehicle load factor (passengers/metre of vehicle length)

N_v = number of vehicles in a train

Equation 2.1 shows that for a given train capacity, line capacity is inversely proportional to the line headway. That is, shorter headways lead to higher capacities.

To meet a system capacity C with a given transit vehicle and one-way route length S_L, n_v vehicles are required. The value n_v is given by Equation 2.2:

$$n_v = \frac{1}{3600} \quad \frac{CS_L}{L_v K_t V_a} \tag{2.2}$$

where V_a = average velocity of vehicles (m/s)

The number of vehicles required to meet a particular capacity is inversely proportional to the average speed. Higher average speeds result in fewer vehicles being required and shorter trip times. For a given vehicle acceleration and deceleration and a given top speed, the average speed in urban rapid transit systems is usually limited by the mean interstation spacing and station dwell time.

The percentage of track which must be occupied by trains is:

$$D_T = \frac{C}{36K_t V_a} \tag{2.3}$$

$$D_T = \frac{100L}{HV_a} \qquad (2.4)$$

where L = train length (m)

D_T may be thought of as "train density". This is a useful concept as it relates the system parameters of capacity, average velocity, and vehicle capacity and it enhances the visualization of possible trade-offs in headway and train length. Figure 1 illustrates three of the numerous combinations of headway and train lengths which will yield a fixed D_T, in this case 30%. All three systems depicted in Figure 1 operate at the same velocity and have the same capacity, yet each has unique characteristics which seriously impact on the overall system design.

Figure 2 shows the passenger capacities that may be achieved through a range of headways and train lengths.

In practice, limitations are placed on the minimum operating headway that is achievable. Such limitations are imposed by the requirement to maintain safe separations on the line and by the time required to get trains in and out of stations. These limits are examined in the following section (2.4).

Two factors also limit the maximum headway at which systems may operate. One limit is dictated by the requirement to maintain a certain level of service or train frequency. A second limit is posed by the need to keep station size and station cost down. If capacity is to be maintained, longer headways lead to longer trains and thus to longer stations.

2.4 Headway Considerations

Headway is a measure of nose-to-nose separation between consecutive vehicles or trains of vehicles. This separation may be measured in terms of distance,

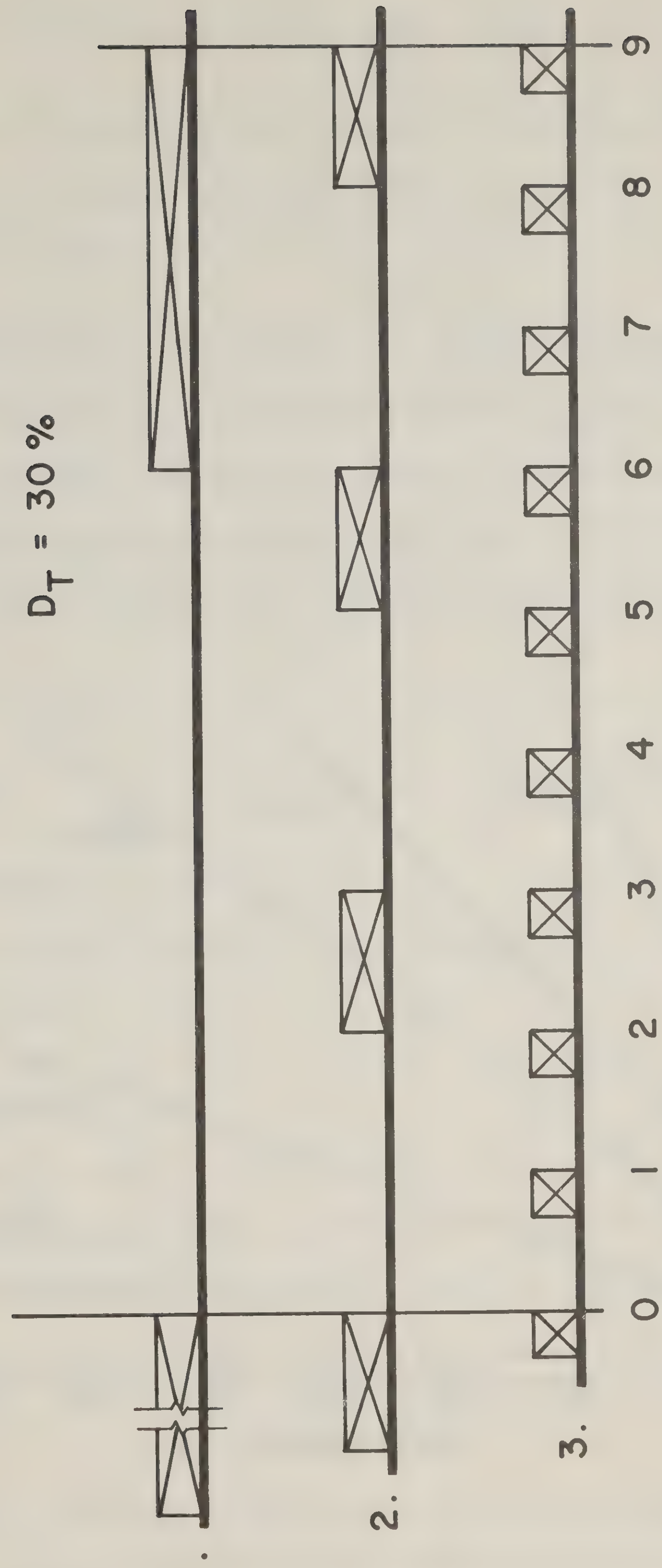


Figure 1; Train Separations for Constant Capacity at Velocity V

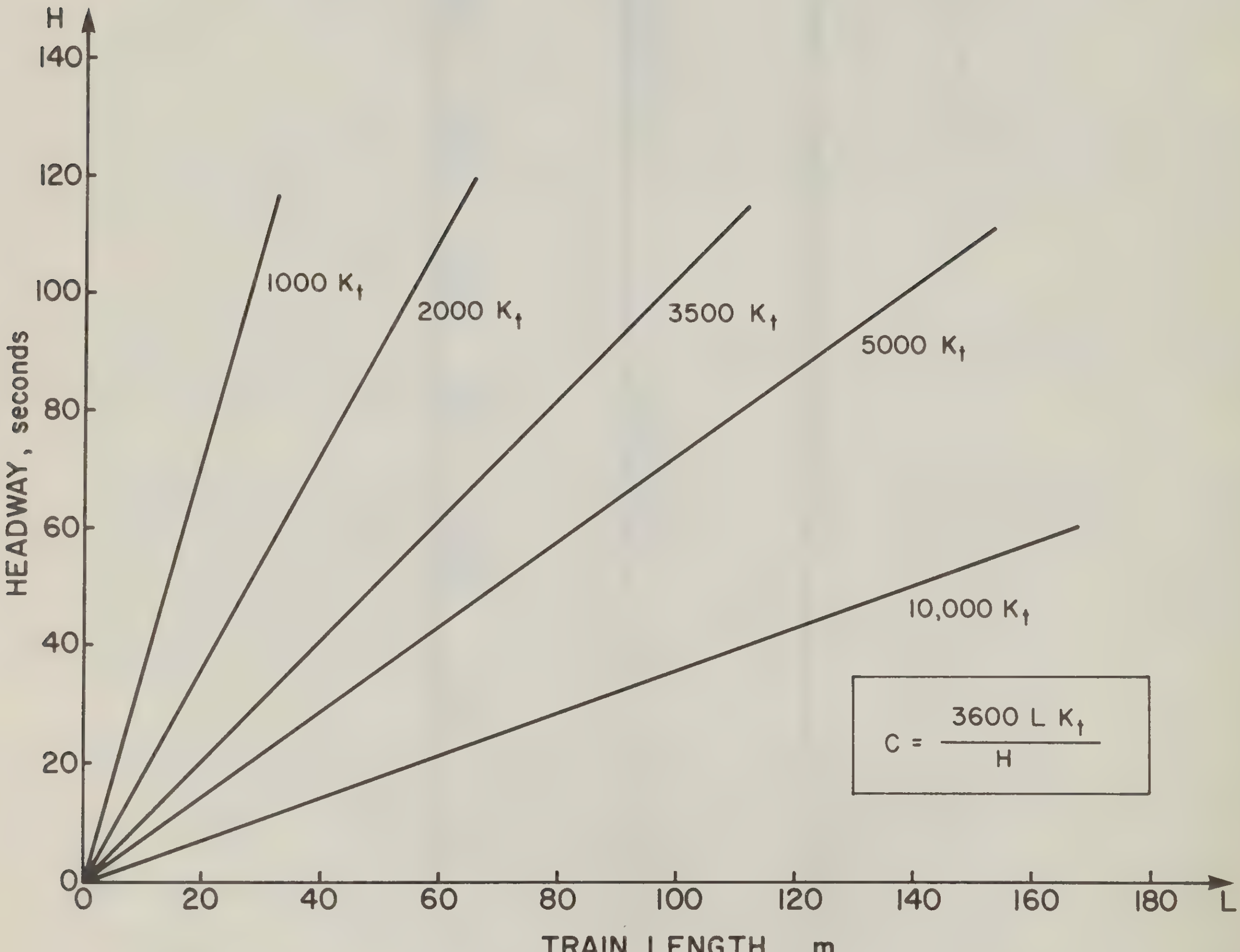


Figure 2; Headway vs. Train Length for Various Capacities

called "distance headway" or in terms of time, called "time headway".

As time headway is inversely proportional to the passenger carrying capacity of a system, it is of prime importance. "Headway" will mean "time headway" in the remainder of the report.

Headway may be measured at any point in the system. For smooth, steady-state operation, the measured headway is the same at any point in the system. The system operating headway is, therefore, dictated by the largest of the minimum achievable headways in the different parts of the system. The various operating areas to be examined to determine achievable headways are:

1/ Stations

2/ Line

3/ Turn-arounds

2.4.1 Station Headway

Station headway is composed of three elements:

1/ Deceleration time

2/ Dwell time

3/ Clear time

(The following paragraphs contain a qualitative description of the three elements, however, a detailed mathematical treatment may be found in Appendix A. Table 1 (page 10) shows achievable station headways as calculated in Appendix A.)

"Deceleration time" consists of two phases: the approach and the deceleration. The deceleration phase is the time to slow the vehicle from line velocity to a stop in the station. The deceleration time is, therefore, determined by the line velocity and the deceleration rate.

Table 1; Line and Station Headways

L (m)	V (m/s)	Moving Block: Station Headway (Figure A17)	Moving Block: Line Headway (Equation 2.8)	Conventional 3-Block: Station Headway (Figure A16)	Conventional 3-Block: Line Headway (Equation 2.7)
15	10	29.5	8.1	42.5	19.4
	20	38.0	13.6	67.0	35.5
	30	46.0	18.4	92.0	52.1
50	10	33.5	11.6	46.0	22.9
	20	42.0	14.8	69.0	37.3
	30	50.0	19.6	93.0	53.3
100	10	38.5	16.6	51.0	27.9
	20	45.5	17.3	71.5	39.8
	30	54.0	21.2	95.0	55.0
150	10	44.0	21.6	56.0	32.9
	20	48.5	19.8	74.0	42.3
	30	57.0	22.9	96.5	56.6

The "approach time" is the period between the time a vehicle is given clearance to enter the station and the time it begins its deceleration. The approach phase is strongly system dependant. As the sophistication and resolution of the signalling system increases, the approach time decreases.

"Station clear time" represents the time required by vehicles to accelerate from a stop and clear the station platform. This time is a function of the train length, the acceleration rate and the line velocity.

"Dwell time" is the time that elapses from the time a train comes to a stop at the station to the time it starts to accelerate from the station. This time includes the door opening and closing times and the passenger loading and unloading times. Dwell time varies with the number of passengers exchanged and this, in turn, varies with the time of day - with the longest dwell times occurring at busy stations during peak travel periods. Assuming that the deceleration time and the clear time are constant for all stations in the system, the station requiring the longest dwell time creates the "bottleneck" in the system. The scheduled system headway must be based on the most constraining station headway in the system to ensure smooth flow throughout the system.

2.4.2. Line Headway

"Line headway" is the time difference between consecutive vehicles passing a point on the line, generally at operating speed.

An examination will show that minimum line headway on systems which meet strict railroad safety standards is always well below the minimum achievable on-line station headways at speeds below 20 m/s as shown in Table 1.

The minimum line headway that may be achieved in a system depends to a large

extent on the safety philosophy adopted during the design phase. The current accepted railroad practice may be stated simply: collisions must not occur under any circumstances. A well-known expression of this philosophy is the "brick-wall" stop concept. This means that a following vehicle must be able to come to a stop behind a vehicle which comes to an instantaneous stop from the operating speed. The following vehicle must, therefore, always maintain a distance behind a leading vehicle equal to one stopping distance plus a reaction distance plus some safety margin under worst case conditions of wind, grade, and deceleration capability. The minimum theoretical line headways are given in somewhat simplified form by the following equations.

$$H_S = k \frac{V^2}{2b} + T_r V + L \tag{2.5}$$

$$H = k \frac{V}{2b} + T_r + \frac{L}{V} \tag{2.6}$$

where V = line velocity (m/s)

H_S = distance headway (m)

T_r = reaction time (s)

k = safety factor

H = time headway (s)

b = deceleration rate (m/s^2)

L = train length (m)

Conventional block* protection systems based on block lengths of 1.35 times the braking distance and 3-block separation between trains yield the following achievable line headways:

$$H_B = 3(1.35 \frac{V}{2b}) + T_r + \frac{L}{V} \tag{2.7}$$

---

*For a definition of block systems see Chapter 5.

Table 1 shows minimum theoretical station headways and line headways employing a moving-block system and achievable line and station headways employing conventional fixed-block systems in use on urban rapid transit systems. These values were obtained using the formulae in Appendix A, Equations 2.7 and 2.8.

$$H_M = 1.35 \frac{V}{2b} + T_r + \frac{L}{V} \qquad (2.8)$$

where H_M = theoretical moving-block line headway for a safety factor of 1.35.

Dwell time is assumed to be 15 seconds. Deceleration rate is 1.2 m/s^2 and reaction time is 1.0 seconds.

Table 1 shows that the station headway for a conventional block control system is greater than the achievable line headway for that system. It is interesting to note that the theoretical minimum station headway achievable with an automated system is still greater than the achievable line headway using conventional block systems, for the conditions given.

As train length and line velocity increase, a point is reached where the line headway exceeds the station headway. In practice, however, this point is not reached on urban rail systems. We can, therefore, conclude that for urban rail systems with on-line stations, the station headway is more limiting than the line headway.

Systems having trunk lines without stations (as shown in Figure 3) represent an exception to the general rule that on-line station headways are more restrictive than line headways in the achievement of minimum headways.

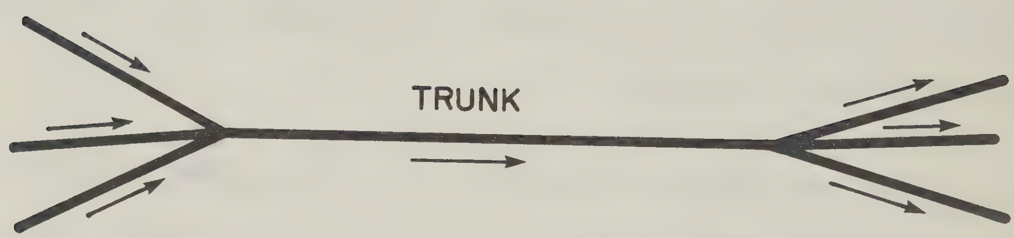


Figure 3; Trunk Line Without Station

In the configuration shown in Figure 3 with vehicles running at the minimum achievable station headways on the three entering lines, the trunk would have to be able to operate at a headway equal to one third the station headway in order to handle the traffic without headway interference. Thus, line headways may be important limitations in some cases.

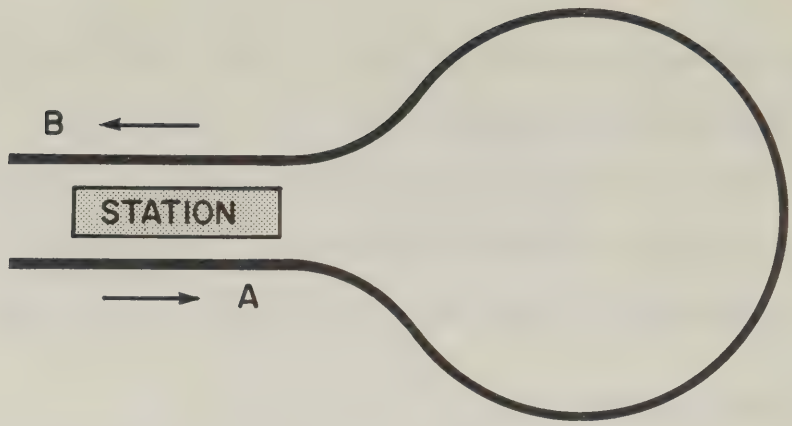
2.4.3 Turn-Arounds

Unless a transit line forms one continuous loop, trains need to change tracks for the return journey at the two ends of the line. In practice, the headways achieved at the turn-arounds are greater than achievable station headways. In order to reduce system headways, therefore, turn-around headways must be reduced. Three common turn-around configurations are shown in Figure 4.

The X turn-around is the most common configuration used in urban transit systems. Track requirements and travel time are minimized while great flexibility of operations is possible. Turn-arounds may be achieved in 1.5 min to 2 min. Headways are limited by certain physical and geometrical constraints as well as by regulations imposed to ensure safety. Efforts to reduce the headway require a detailed examination of both the physical limitations and the operational requirements. Automation may yield reductions in the order of 10 s to 15 s only.

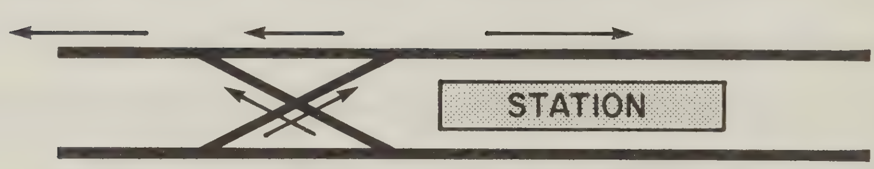
The Y turn-around employed with a conventional 2-block system will allow turn-around headways in the order of 80 s to 90 s for a 100 m train. This is greater than the approximately 20 s practically achievable station headway (see Figure A3) at a top line speed of 20 m/s.

Referring to Figure 5, a rough calculation of turn-around headways may be made.



(a) LOOP TURN-AROUND

(b) Y TURN-AROUND



(c) X TURN-AROUND

Figure 4; Turn-Around Configurations

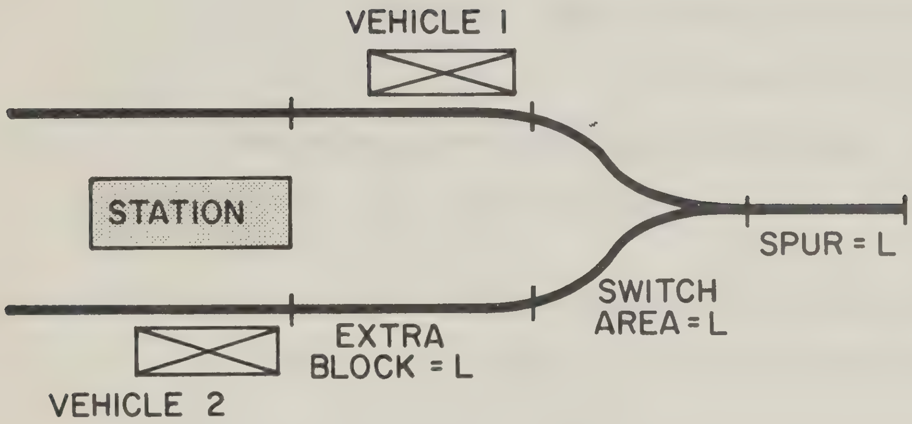


Figure 5; Y Turn-Around, Conventional 2-Block System

Vehicle 2 may leave the station as soon as Vehicle 1 clears the switch area. The turn-around headway is the time required for Vehicle 2 to traverse the switch and arrive at the position shown for Vehicle 1. The distance to be travelled is 5L where "L" is the block length. The minimum block length is equivalent to the train length. The time required to travel 5L including two accelerations and one deceleration is :

$$t_t = \frac{\left(5L - \frac{3V^2}{2b}\right)}{V} + \frac{3V}{b} + 2t_s \tag{2.8}$$

where L = block length = train length (m)

V = top speed (m/s)

b = acceleration and deceleration rate (m/s^2)

t_s = switch confirmation times

The above turn-around time does not include the time required for the train's driver to move from one end of the train to the other.

With a moving-block system, the "Extra Block" (Figure 5) may be eliminated. The turn-around equation then becomes:

$$t^1_t = \frac{\left(4L - \frac{3V^2}{2b}\right)}{V} + \frac{3V}{b} + 2t_s \tag{2.9}$$

Table 2 shows the achievable turn-around headways with a Y turn-around. The top speed reached at the turn-around is assumed to be 10 m/s. The switch confirmation time is 10 s, and the deceleration/acceleration is equal to $1\ m/s^2$.

Table 2 shows that automation results in a saving of 3 s to 15 s in turn-around headways, depending on the train length. Reference to Table 1 (page 10) reveals that the turn-around headway is less than the station headway for low

Table 2; Turn-Around Headways for Y-Type Turn-Around

Train Length (m)	Conventional 2-block System	Moving Block System (s)
15	42	39
50	60	55
100	85	75
150	110	95

station approach speeds and very short trains. At higher station approach speeds and larger train lengths, the turn-around headway is limiting.

Switch confirmation times contribute significantly to the turn-around headway for the Y configuration. Where turn-around headways are critical, serious study should be given to possible switch improvements.

The loop turn-around offers great potential for systems requiring short headways. As no geometrical limitations exist, headways as short as station headways may be achieved. In effect, the end station may simply be considered as two stations. One station is entered as the train arrives at the end station and the second as the train completes the loop. Each of these dual stations may be transacted in one station headway time, with no bottleneck being created.

The loop is the simplest configuration, requiring no switches or additional signalling. Vehicles need not change directions and the driver does not change his position for the return journey. However, the configuration is costly in terms of the large area of land enclosed by the loop as well as the requirement for extra length of track. An 80 m radius loop encloses over 20 000 m^2 (about 5 acres) and requires more than 500 m of track right-of-way. The loop configuration results in a shorter end station turn-around headway than the X or Y turn-arounds and thus yields shorter achievable system headways. The reduction in the achievable headway may be used to reduce train lengths, to improve service frequency at constant capacity, or to increase capacity if the train length is maintained.

Any increase in system capacity brought about by reducing headways brings about an increase in the fleet size. The increase in fleet size is directly

proportional to the increase in capacity.

Reducing the turn-around headway below achievable station headway values once again places the system bottleneck at the stations where station headway depends on the train protection system employed. A moving-block system* deserves serious consideration if headways less than those achievable through the use of fixed-block systems are desired. As stated previously, station headways may be reduced by as much as 40% by introducing a moving-block protection system. The station headway may also be reduced by a number of other techniques such as shorter blocks near the station or cab signalling.

Loop turn-arounds are a technique which existing systems may incorporate into their structure to quickly and relatively inexpensively increase capacity. If the system headway is limited by the station headway, the introduction of automation to reduce station headway and increase capacity may be cost-effective. Chapter 7 considers the cost and benefits of automation.

2.4.4 Headway Dynamics

In the preceding analysis some simplifications have been made. One of these is the consideration of steady-state operating conditions only. In a true operation situation, however, the relative position of vehicles and their headways are changing constantly.

Ideally, when no delays are encountered, theoretically calculated schedules and headways can be met. In practice, however, one may not design systems to operate at minimum theoretically achievable headways. Any disturbance or

---

*For a definition of block systems see Chapter 5.

delay to a vehicle will cause the disturbance to travel back up the line causing vehicles to slow down or halt on the line. In order to avoid excessive delays that may be caused by such "traffic jams", headways need to be extended somewhat from the ideal minimum. Therefore, the achievable operating headway is the minimum theoretically achievable headway plus a marginal headway.

To obtain a more realistic estimate of achievable operating headways, computer simulation of the system is required. Through simulation, the dynamic effects of delays and disturbances on the overall system may be predicted. Headways are then extended so that intermittent but unavoidable delays at stations, or delays to particular vehicles, will not affect the system adversely. A simulation study of this nature requires detailed information regarding system operating characteristics, station design, operating strategy, and schedules as well as accurate statistics of vehicle and station delays and disturbances.

2.4.5 Choice of Headway

The preceding paragraphs looked at the headways achievable at different points in a system: on the line, in stations, and at the turn-arounds. The operating headway chosen for a system must be greater than the largest theoretical minimum achievable headway at any point in the system.

Figures 6 and 7 show the theoretical minimum achievable headways for various parts of the system superimposed on a graph relating train length, headway and system capacity. Figure 8 illustrates the way Figures 6 and 7 are to be interpreted. Figure 6 illustrates minimum headways for moving-block systems while Figure 7 gives headways for conventional block systems.

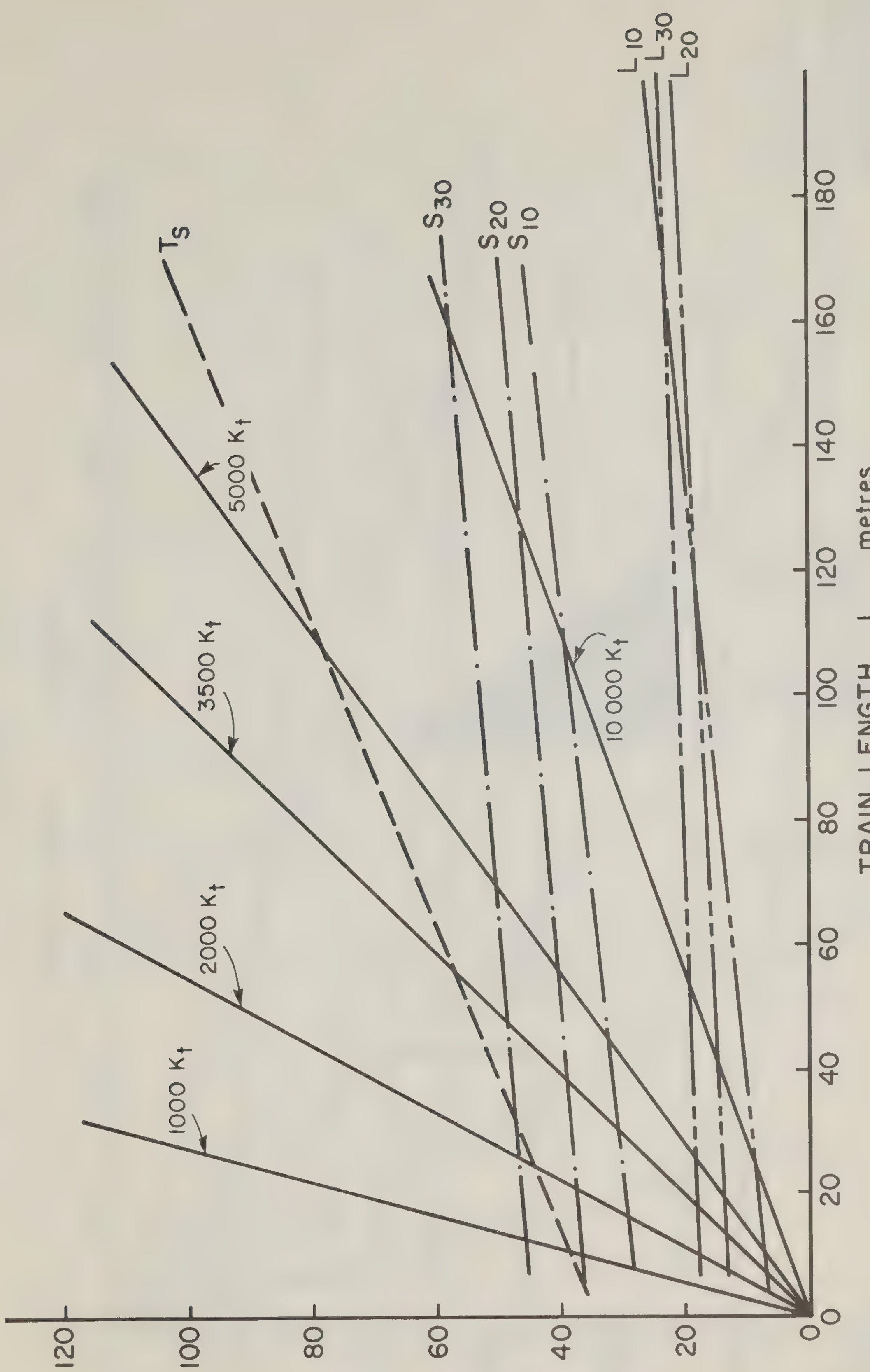


Figure 6; Minimum Achievable Headways for Moving-Block System

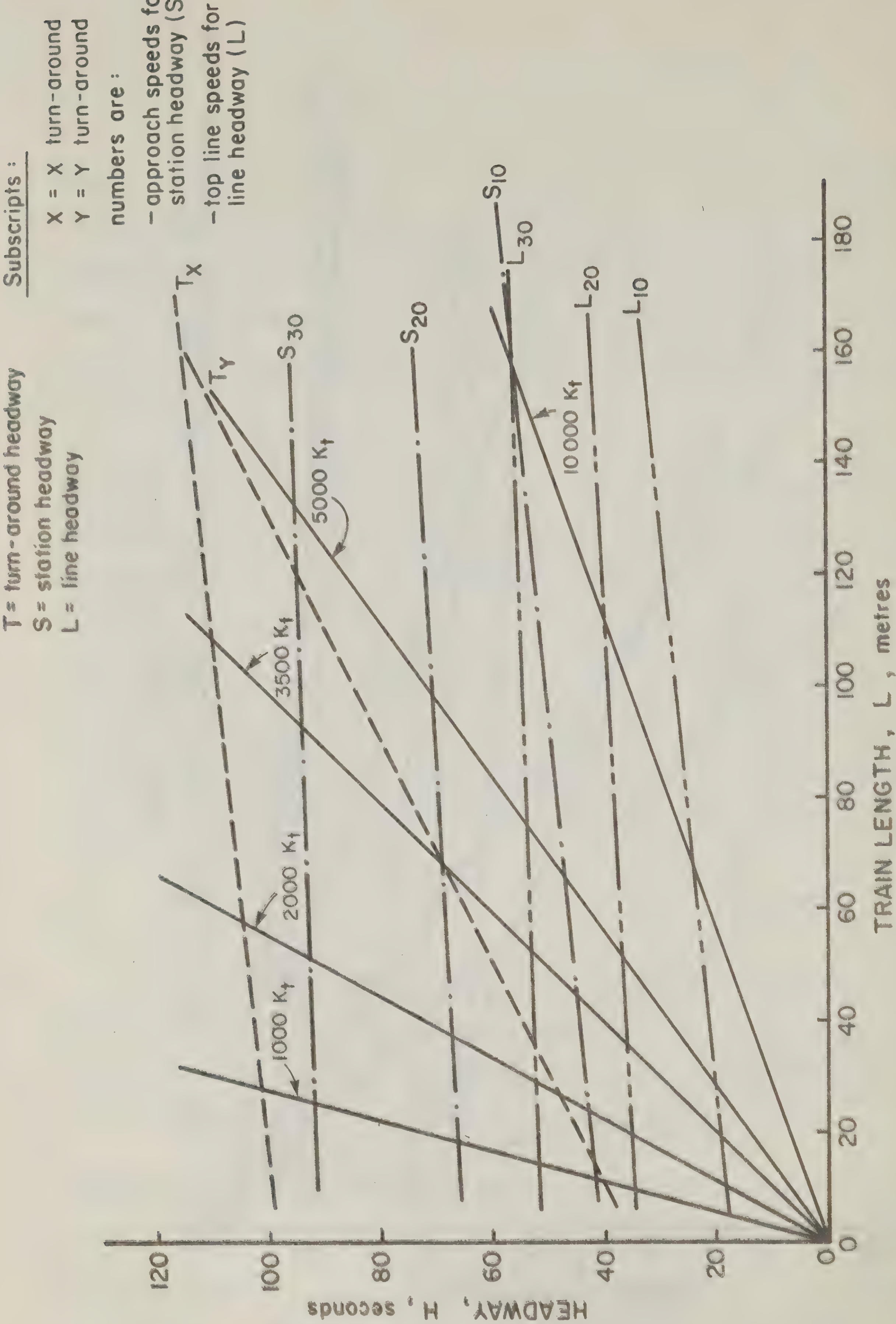


Figure 7; Minimum Achievable Headways for Conventional Fixed-Block Systems

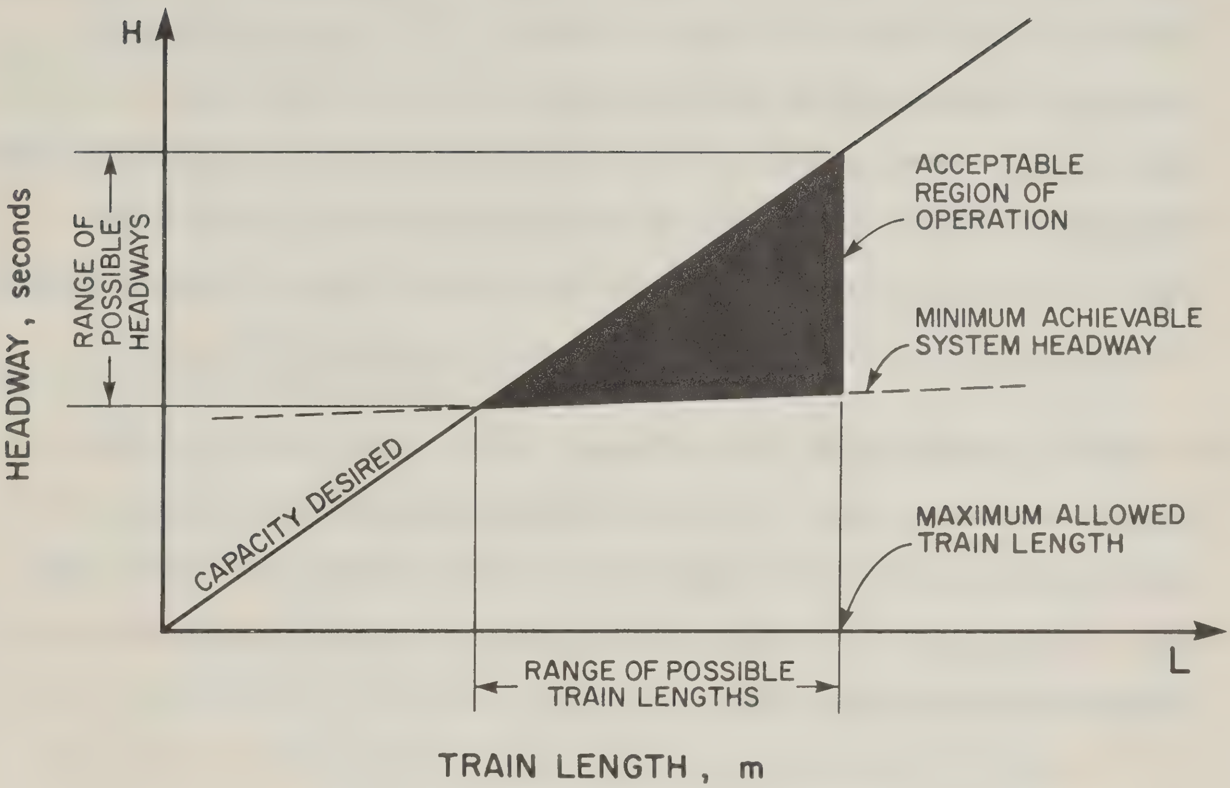


Figure 8; Operational Headway

An upper limitation on train length is posed by the desired station configuration. As the capacity/headway/train length relationship shows, train length increases with headways for fixed system passenger capacities. Available space and construction costs will dictate the maximum acceptable train length.

Between the two extremes of headway dictated by the minimum achievable headway and maximum desired station platform length, a wide range of possible headways still exist to meet the desired capacity. Level of service is traded off against the cost of maintaining that level of service. With manned trains, operating staff costs decrease as headways increase. Better service frequencies, however, result in an increase in ridership and revenue.

In general, headways vary with the time of day. During peak travel periods the system operates close to the minimum achievable headway. During off-peak hours, headways are increased and trains may be shortened. The final choice of the operating headway must be based on a detailed analysis of the cost and benefit to the system.

3. AUTOMATIC TRAIN CONTROL (ATC) FUNCTIONAL REQUIREMENTS

Automatic Train Control comprises three subsystems:

1/ Automatic Train Protection

2/ Automatic Train Operation

3/ Automatic Train Supervision

The functional requirements of each of the above three subsystems are defined below.

3.1 Automatic Train Protection (ATP)

The ATP System is responsible for train safety. It is required to protect trains from collision, sideswipe, and derailment.

3.3.1. Failsafe Design Criteria

Both trainborne and wayside equipment which makes up the ATP system must meet the following failsafe design criteria:

a/ Self-detecting component failures will cause the train to stop or run at reduced speed.

b/ Non self-detecting component failures will not cause unsafe consequences and will not, when added to other failures, cause unsafe consequences.

c/ Any number of simultaneous component failures attributable to the same cause or related causes will not cause unsafe consequences.

3.1.2. Train Detection and Track Surveillance

a/ Trains must be detected in both the interlocking and non interlocking areas of the track network.

b/ The intended route, identity and general condition of all trains in the network must be determined at all times.

c/ The general condition of the entire track network as well as the position of

all switches are determined throughout the operating day.

3.1.3. Train Separation

a/ Consecutive trains must be kept at a separation of at least one safe braking distance to be determined by the train separation system. The safe braking distance of a train takes into account its train length, loading condition, velocity, deceleration, jerk limit, position along the track, and the track condition.

b/ The separation system calculates and determines the safe velocity for a train to follow in order to maintain safe separation from its preceeding train, the position of which is determined by the train detection system.

3.1.4. Interlockings

a/ Interlockings are provided at junctions, crossover, and other critical locations as required.

b/ The function of interlockings is to establish and maintain safe train routes.

c/ The interlocking system provides route security along the track and through switches by calculating and commanding safe switch locking to disallow conflicting train movements.

d/ At grade crossings, interlocking is required to forbid conflicting movement between train movements and road traffic.

e/ The switch locking and movement is continuously monitored and checked for integrity through the train detection and surveillance system. Remedial emergency action must be taken in the event of nonconformance.

3.1.5. Train Overspeed Protection

The trainborne overspeed protection equipment protects a train from unsafe overspeed conditions by:

a/ Measuring the actual speed of the train.

b/ Continuously (or periodically) comparing the actual speed with the most restrictive command speed (which is the lower of the command speed which has taken into account the civil speed limit and the safe speed derived from the separation system).

c/ Applying the emergency brake if the actual speed exceeds the most restrictive command speed by a prescribed amount.

d/ Having absolute stop assurance.

3.2 Automatic Train Operation (ATO)

The Automatic Train Operation (ATO) system performs the on-board functions traditionally performed by the motorman and conductor. It regulates train speed, controls programmed stop, controls door operation and dwell time, and initiates starting. Each of these functions is defined below.

3.2.1 Automatic Speed Regulation

a/ The automatic speed regulation receives and acknowledges speed commands from Automatic Train Supervision (ATS).

b/ The automatic speed regulation system regulates the train speed to the command speed from ATS approved by the automatic train protection system.

c/ The velocity regulation system must be acceleration- as well as jerk-limited to conform with the passenger safety and comfort criteria.

d/ Central Control shall automatically be informed as soon as an underspeed condition is detected.

e/ In case no speed command is received, the speed regulation system takes it as a zero speed command.

3.2.2. Programmed Stopping

The programmed stopping system causes the train to decelerate to a

stop at a passenger station platform with a prescribed stopping accuracy by:

a/ Determining the relative distance between the train and the stopping point at the approach station at which programmed stopping is initiated.

b/ Knowing a velocity distance profile for the train to follow to an accurate station stop when programmed stopping is initiated.

c/ Regulating the trajectory of the train to the velocity distance profile using the automatic speed regulation system.

It has the added capability of

d/ Causing the train to skip any station stop by inhibiting the programmed stop function.

3.2.3 Door and Dwell Control

The door and dwell control system controls the operation of the train doors and the length of time that a train remains in a station by performing the following functions:

a/ Prohibiting door opening while the train is in motion.

b/ Opening the doors adjacent to the platform only when the train is completely stopped and correctly berthed at the platform with full service brakes applied.

c/ Initiating station dwell timer once the train doors are open at the station.

d/ Commanding door closure at the expiration of the dwell time specified by ATS.

e/ Verification of door closure status which also indicates that no passengers or their possessions are caught between doors which would remain open otherwise.

f/ In case doors still remain open upon verification, commanding door closure again. If door closure status cannot be obtained after a specified number of trials, station personnel must be alerted.

g/ Upon confirmation of door closure status, the brake is released and the train is handed over to the automatic speed regulation system and ready to start.

3.3 Automatic Train Supervision (ATS)

The Automatic Train Supervision (ATS) system supervises scheduling and monitors the operation of the network by performing some, or all of the following functions:

a/ Dispatching trains to revenue service.

b/ Determining a reference plan which contains daily nominal train routes and schedule based on demand analysis, simulation and past experience. The train schedule contains the nominal velocity, position, station dwell times and identity of all trains operating in the network as a function of time. The nominal train route table contains the intended routes and switches as a function of time for all trains.

c/ Determining from the train detection and surveillance system, the actual position, velocity, and identity of all trains operating in the network and the position of all switches at all times.

d/ Comparing the actual performance of all trains with the reference plan.

e/ Maintaining schedule or minimizing delays by adjusting the performance of some or all trains if a discrepancy exists between the actual performance and the reference plan due to minor disturbances.

f/ Revising schedule if performance modifications of trains are not sufficient to restore the system to stable operations.

g/ Displaying the entire system's status showing the progress of trains in the network, the support facilities and the electrification system.

h/ Reacting to equipment failures, abnormal and emergency situations with minimum inconvenience to passengers.

i/ Logging system data such as past system performance to aid in update of fixed-time or demand-responsive scheduling, system component failures for ease of maintenance, number of car-miles for each train, number of trains or cars in the network including those in the maintenance area, etc.

j/ Removing trains from revenue service.

3.4 Summary

Table 3 provides a summary of the key functions of Automatic Train Control. The inter-relationship of ATP, ATO, and ATS are shown in block diagram form in Figure 9.

Table 3; Summary of ATC Functions

Automatic Train Protection (ATP)	Automatic Train Operation (ATO)	Automatic Train Supervision (ATS)
	-Automatic speed regulation	-Scheduling and monitoring network operation
-Train detection and track surveillance	-Programmed stopping	-Performance modifications
-Train Separation		
-Interlockings	-Door and dwell control	-Reacting to abnormal and emergency situations
-Train overspeed Protection		-Displaying system status
		- Logging System data

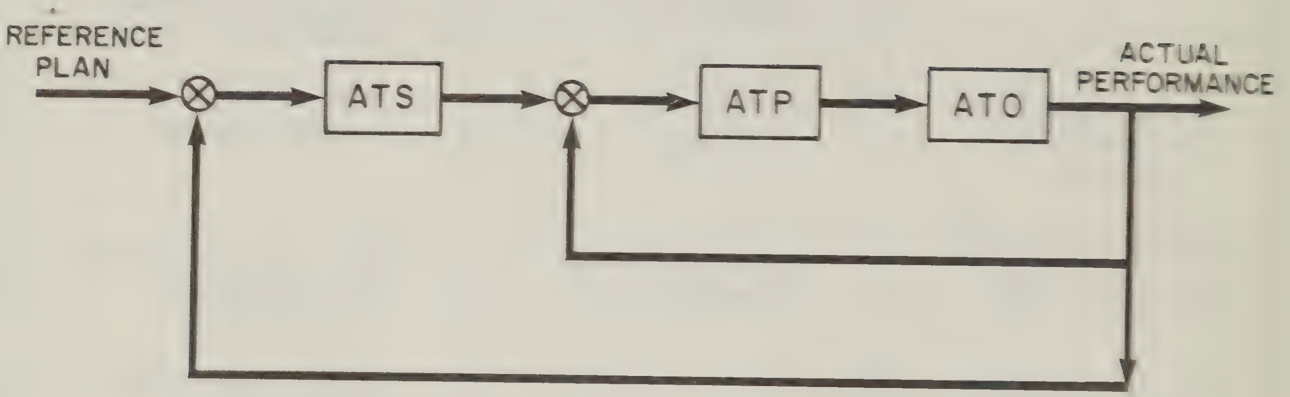


Figure 9; Information Flow Block Diagram for ATP, ATO, and ATS

4. SIGNALLING REQUIREMENTS

4.1 Introduction

The successful evolution of transportation systems towards higher operating speeds and efficiencies depends to a large extent on the provision of additional information to the driver and/or on the automation of certain functions which the operator is ill suited to perform.

The requirement of additional information and the automation of certain functions is based primarily on safety considerations arising from the psychological and physical limitations of human operators. The importance of the economic aspects of signalling was recognised early but only in recent times did it assume prime importance.

Many system-specific factors contribute to the definition of human operator performance limits for a given driving task. Ideal limits of line-of-sight, ranges of driver reaction times, etc., can be readily derived. For a specific system, however, track geometry, exclusiveness of right-of-way, environment, operational strategy, etc., all contribute to the definition of the actual limits beyond which the human operator requires supplemental information to perform his task safely and efficiently.

It is well known that human operators are ill suited to perform routine tasks over prolonged time periods. It is for this reason that information in the form of signals must be reinforced by train protection devices which can warn, and, if necessary, override the driver. If the task of safe and efficient driving exceeds the physical and/or psychological performance capabilities of the human operator (e.g. high speed driving), signalling must be used to supplement and extend his capabilities. Psychological and physical considerations very often overlap in the sense that both physical and

psychological reasons may dictate the need for additional information and that physical factors often have psychological correlates.

4.2 Physical Limitations of Human Operators

4.2.1 Limitations of Line-of-Sight Driving

Line-of-sight driving requires that the sighting distance at all times, for all operating speeds and conditions exceed the worst case braking distance (including driver reaction time). Since the sighting distance varies with prevailing visibility conditions and since the braking distance is a function of weather conditions, line-of-sight operations can never be very safe or reliable as collisions may occur. Nevertheless, it is worthwhile to examine the performance of a human operator in a line-of-sight operation and try to establish an upper limit beyond which line-of-sight driving becomes 'totally' unsafe.

The ability of a driver to follow a preceding vehicle can be examined in the following way. Assuming that the leading vehicle is always visible and originally at least one safe braking distance ahead of the vehicle following it, the driver of the following vehicle estimates the angle subtended by the leading vehicle and if, at some later time, this angle has increased sufficiently that a change in the subtended angle can be detected, the operator of the following vehicle concludes that he is closing up on the leading vehicle. Thus, the change in the size of the image on the retina of the eye is the psychological correlate of the physical process of estimating the variation in separation between two vehicles[1].

---

[1] Michaels, R.M. and Cozan, L.W., Perceptual and Field Factors in Car Following, Highway Research Board, Bureau of Public Roads, Washington, D.C., 1963.

Figure 10 shows various threshold values of detectable change in subtended angle at different speeds. Safe braking distances for vehicle deceleration rates of 0.8 m/s^2, 1.2 m/s^2, and 2.0 m/s^2 are plotted on the same graph. The graph shows that for any given deceleration rate, even under good visibility conditions, the driver's capabilities are exceeded above certain speeds. Simply stated, at high speeds the required braking distance increases and before the driver perceives a change in subtended angle, the following vehicle may be too close to avoid a rear end collision.

Figure 11 shows the "safe" and "unsafe" velocity deceleration regions. In the safe regions, under good visibility conditions the drivers are within their physical performance capabilities (based on sight only) to operate the transit vehicle. Outside the safe regions, the physical capabilities of the driver are exceeded and line-of-sight driving is unsafe.

Table 4 briefly summarizes the results which show that for transit operations with guaranteed deceleration rates of 0.8 m/s^2 to 1.2 m/s^2, the physical limits of human operators for line-of-sight driving are exceeded for operating speeds of approximately 20 m/s and above. (A note of caution: this should not be taken to imply that for speeds of 20 m/s, line-of-sight driving is safe and efficient.) The fallibility of the human operator must always be considered and this consideration may dictate the use of signalling and forms of automatic train protection at much lower speeds.

4.2.2 Limitations on Wayside Visual Signalling

In the previous section (4.2.1) it was shown that line-of-sight driving at high speeds and for the given deceleration rates was unsafe, since the physical capabilities of the operator to detect separation were exceeded. In the following, it will be shown that even if wayside visual signalling is

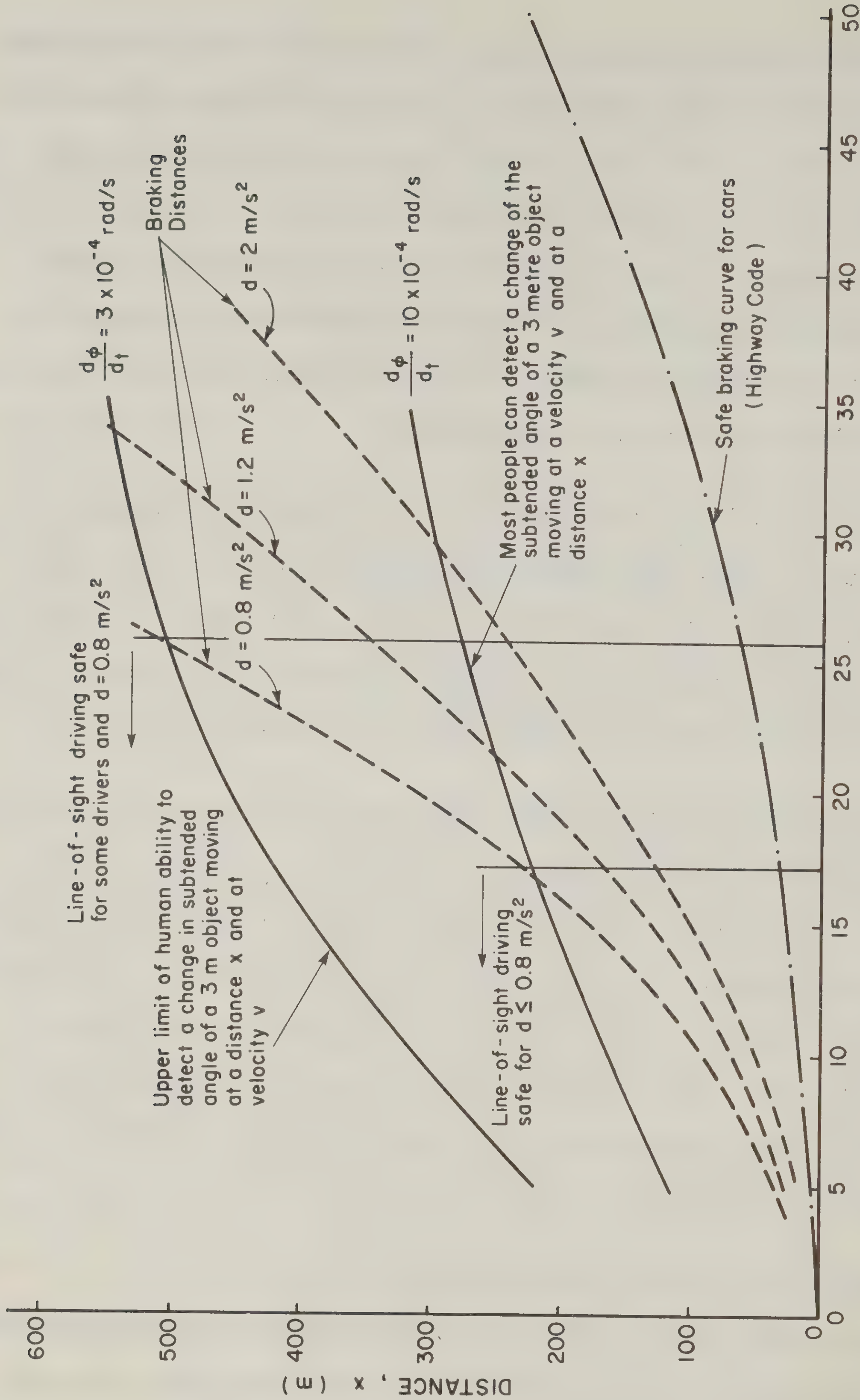


Figure 10; Constraints on Line-of-Sight Driving

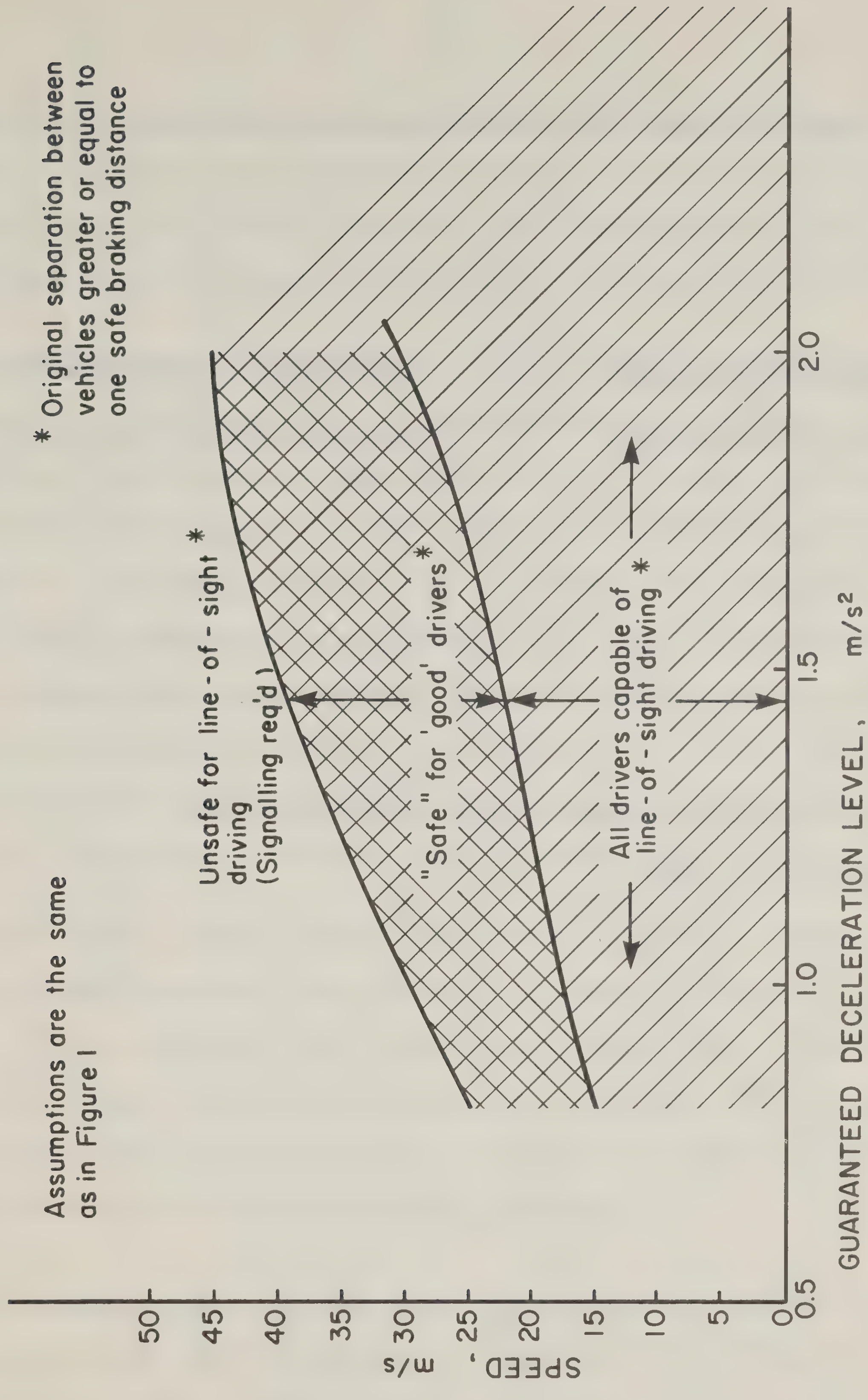


Figure 11; Limits of Line-of-Sight Operation

Table 4; Limits on Visual Driving Without Signalling

GUARANTEED DECELERATION m/s^2	MAXIMUM SPEED WITHOUT SIGNALLING m/s^2
0.8	17.5
1.2	22.0
2.0**	30.0*

* Streetcars usually operate at speeds much below 20 m/s.

** Magnetic Track Brakes

used to supplement the driver's capabilities, high speed driving may overtax both his physical and psychological capacities. To permit operation under these conditions, a form of signalling which is independent of the sighting distance (e.g. cab signalling) and automatic train protection is necessary.

For a fixed sighting distance (200 m), Table 5 shows the sighting interval, (the time that the signal is visible to the driver proceeding at a given speed) and the frequency at which the driver has to inspect the wayside to make sure that he does not miss the next signal. To be certain that the signal is not missed, the driver has to sample at least twice within a sighting interval. With increasing speed, the task of trying to observe the next signal places a high demand on the driver. As the speed is increased, track profile variations, signal placements etc., are closer together and the frequency of operations which a driver has to carry out rapidly increase to the level where the driver is not sufficiently reliable.

Psychologically, man is limited in his ability to attend to signals coming simultaneously from more than one source. If signals originate from two or more sources, the driver will process them in serial order. With increasing operating speeds and increasing frequency of tasks, the driver requires pre-processed information and, of course, some form of automatic train protection in case he misses or misinterprets the information.

Cab signalling, which in effect puts the wayside signal into the driver's compartment and which can provide a warning signal and fail-safe driver override, is used extensively in high speed applications. As indicated in Figure 12, cab signalling is a legal requirement for trains operating in excess of 50 m/s in the U.S. and in West Germany. For speeds above 70 m/s, operation under human control, even with cab signalling, becomes unsafe and

Table 5; Human Information Processing Capability in Train Operation

VELOCITY m/s	SIGHTING DISTANCE m	SIGHTING INTERVAL s	REQUIRED SAMPLING FREQUENCY
10	200	20	once every 10 s
20	200	10	once every 5 s
30	200	6.7	once every 3.85 s
40	200	5	once every 2.5 s
50	200	4	once every 2.0 s
60	200	3.3	once every 1.65 s
70	200	2.9	once every 1.45 s

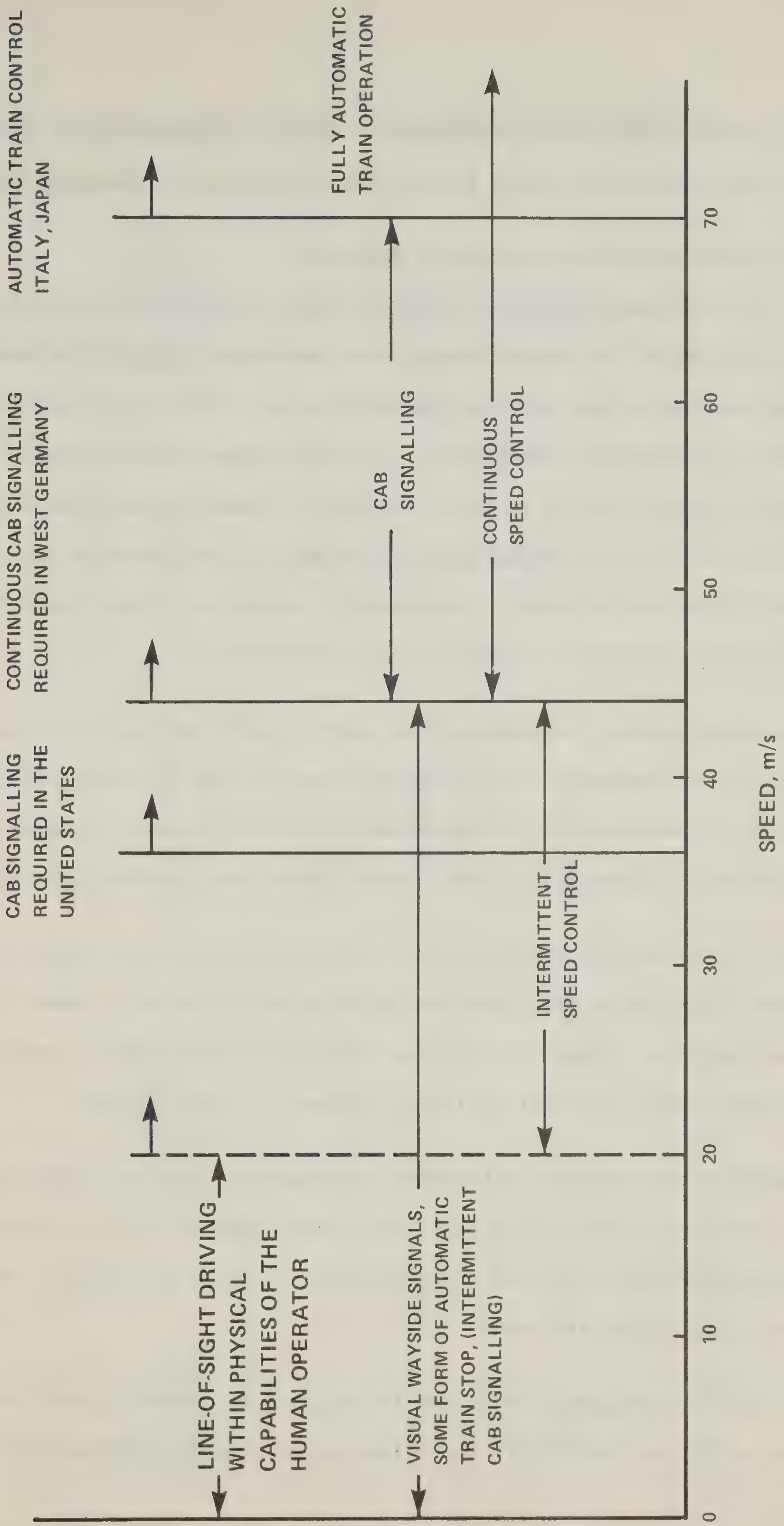


Figure 12; Signalling vs. Speed

full automatic operation is necessary. Drivers are normally retained, however, for track surveillance, door control, and assistance in abnormal situations.

4.2.3 Limitations on Short Headway Operation

For very short headways, in the order of few seconds or shorter, the reaction time of the driver becomes very important. Table 6 relates the rate of decision making and information handling as a function of multiple choice between equiprobable alternatives. The table shows that in dynamic situations calling for high rate of decision making, the human operator can handle only a limited selection of responses. His reaction time increases rapidly as the number of choices increase. Consequently, operation at very short headways requires full automation without human interference.

For headways normally encountered in rapid transit systems (60 s and above) signalling is necessary for interlocking and for the safe and punctual movement of vehicles. A very brief analysis will be used to support this claim. The analysis is intended to show trends rather than produce absolute values.

Table 7 is a very simplified list of the functions which a driver has to perform in driving a train and the decisions which he has to make. Table 7 is not complete. There are more functions to be performed and more decisions to be made. For the brief analysis, however, it will suffice.

In Table 8, the frequency with which each function has to be performed is shown as a function of speed. The last two columns show the rate of decision-making required from the driver and the percentage to which the driver's decision-making capabilities are taxed.

As the speed increases, the driver is required to operate closer and closer to his capabilities. At 30 m/s the driver is required to operate at 54% of his

Table 6; Rate of Decision Making and Information Handling*

Number of Equiprobable Alternatives	Uncertainty in Each Decision (bits)	Reaction Time (s)	Decisions/s	Max. Rate of Handling Information bits/s
1	0	0.187	5.35	0
2	1	0.316	3.16	3.16
5	2.32	0.487	2.05	4.76
10	3.32	0.622	1.62	5.38
32	5	0.845	1.184	5.92
64	6	0.972	1.03	6.18
128	7	1.105	0.91	6.34

*From: "The Analysis of Skills in Driving", Australian Road Research, March 1964.

Table 7; Train Operation Functions and Some of Their Information Contents

Function	Action	Number of Decisions
R.O.W. Surveillance	Go/Stop	2
Signal Observance	Go/Stop	2
Speed Observance	Accelerate/ Decelerate	2
Instrument Surveillance	Go/NoGo	2

Table 8; Comparison of Driver Decision Making Capabilities and Requirements

Instrument Obs. (s /event)	Speed (m/s)	Signal Surveillance* (s /event)	ROW Surveillance** (s /event)	Speed Observ. (s /event)	Events/s	Req'd Decisions/Possible Decisions
5	10	10	1	5	1.4	44%
5	20	5	1	5	1.6	50%
5	30	3.85	1	5	1.7	54%
5	40	2.5	1	5	1.8	57%
5	50	1.65	1	5	2.1	67%

* From Table 5

** Based on the assumption that a driver must react within 1 s to a right-of-way violation occurring in his line of vision.

rate of decision-making capabilities; at 50 m/s he is already at 67%.

Filmed records of driver's eye movements (while driving an automobile) indicate that the maximum rate of sampling from separate sources is about 1/s to 1.4/s. A glance of 0.75 s should be enough to give a reasonable estimate of the R.O.W. status and the signal status. Roughly 1 s is required for speed checking and 1 s to scan the instrument panel. As the second last column in Table 8 shows, the number of information sources requiring attention prevents the allocation of sufficient time for reliable judgements. The driver is forced to fall back on experience and this can become dangerous - particularly at high speed.

Close examination of Figure 10 also reveals that human operators without the aid of supplemental information (e.g., signals), are ill suited for headway keeping in rapid transit applications. At a nominal time headway of 60 s and a speed of 20 m/s, the trains are separated by 1200 m. Figure 10 shows that the human operator cannot detect any change in the separation between trains at this speed and this nominal separation. In fact, the trains have to be within roughly 450 m (time headway of 22.5 s) before the operator of the following train can detect a change in the headway.

Interestingly, the human operator appears to be better equipped to keep short headways. However, at short headways the unreliability of the operator in a continuous control function necessitates signals and overrides. The 'open-loop' approach, in which all the trains follow a predetermined pattern, breaks down as soon as delays or disturbances occur. In fact, at low speeds (and with the current state-of-the-art) driver control may provide the shortest headway. As shown in Figure 10, automobiles, which have higher deceleration capabilities, can operate safely at higher speeds and shorter headways than

trains. For autos, safe braking distance separation is not always a requirement with driver operation because drivers can anticipate more than one vehicle ahead and because they have some capability for lateral manoeuvres.

4.2.4 Other Limitations

One important consideration in operating under visual wayside signalling is the requirement of being able to reduce vehicle speed to the next most restrictive speed within one sighting distance. If signal repeaters are used, or if additional signal aspects are introduced to effect a more gradual speed transition, the headway is increased by the insertion of additional sighting distances.

4.3 Economic Aspects of Signalling

The purpose of signals is the transmission of information to employees in charge of the operation of trains.

Safety was undoubtedly the original purpose for which signals were installed and the fundamental principle of signals was centered around this original purpose. In actual application, however, it is difficult, if not impossible, to separate the principles pertaining to safety from those pertaining to facility of train movement. Therefore, one must always keep in mind that the introduction of signalling will alter the economy of the system. To prepare a realistic economic statement, all of the following must be considered:

1/ Operating methods

2/ Change in elements of annual cost due to signalling

3/ Value of each element to the system

4/ Actual change in annual cost (2/ x 3/)

5/ Acquisition, installation, and maintenance costs of signalling

6/ Return on investment

All of the above are specific to a system and without investigating a specific system, only general statements can be made. In later sections, however, simple cost analyses will be conducted to indicate the merits of various automation levels.

In general terms, signalling affects all three aspects of transit organization: the permanent way, the rolling stock, and the operating method.

Operating method is affected since signalling permits a higher system capacity, more punctual service, and service independent of environmental conditions. Rolling stock is affected since signalling permits higher schedule speeds which results in a smaller vehicle fleet. Permanent way construction, geometry, repair, and maintenance are also affected to some extent by signalling. This list is by no means exhaustive, since all operating cost elements could be shown to be affected by signalling. It is only meant to show that signalling has a strong economic impact and that under some conditions the economic aspect of signalling will outweigh the safety considerations.

4.4 Summary

The requirement for additional information to the operator has been discussed. Qualitatively, it was shown that speed and operating headways may set requirements on the human operator which can exceed his physical and psychological capabilities. Even if the normal physical and psychological limits are not exceeded, the human operator is prone to making mistakes - particularly in routine tasks. This unreliability of the human operator dictates the use of signalling and safety devices even at low speeds and relatively long headways. As shown in Figure 12, higher speeds require a higher degree of automation whereas at lower speeds, additional information to the driver and safety override devices are satisfactory from a safety point-of-view.

There is also a change in the way information is processed and provided at higher speeds. Speed control becomes more important at higher speeds.

Whereas this chapter is intended to provide a general overview of why various levels of signalling and train protection are necessary, the following chapters will provide an insight into the hardware organization of such devices and their typical applications. Economic aspects of signalling were introduced in general terms. More detailed analysis of signalling and automation economics will be provided in a later section of this report.

5. BLOCK SYSTEMS

Block systems are headway control systems, i.e. systems which deal with the safe separation of all the vehicles on a track. Basically, there are two types of block systems in use: the fixed-block system and the moving-block system. The following paragraphs will discuss the salient features of these two systems.

5.1 Fixed-Block Systems

In fixed-block systems, the track is divided into segments, or blocks, each of which can be occupied by only one train at any given time and consecutive trains are always separated by one or more blocks irrespective of their speed. This approach to train separation is a space interval method.

The entrance and use of a block is governed by the block signal placed at the beginning of the block. This can be a physical wayside signal or an electrical signal indicating the beginning of a block. The various signals (minimum of 2) which are displayed to the driver (or the ATO) are referred to as signal aspects. It is usual practice to require that the block immediately behind an occupied block be a stop block in which a following train must be brought to a stop by the emergency brake. Since, from a safety point-of-view, trains must always be a safe braking distance apart, blocks are usually made equal to integer fractions of the safe braking distance (e.g. 1, 1/2, 1/3, etc).

Calculation of the safe braking distance usually considers the worst case condition and is based on the concept of a "brick wall" stop. Violation of the safe braking distance separation results in the emergency braking of the following vehicle. Even though the safety condition is satisfied if trains follow each other by one safe braking distance only, the division of the track into coarse blocks results in a 1 block uncertainty of train location. The

uncertainty results in false alarm emergency braking whenever the following train enters the stop block (the block immediately behind the one occupied by the preceding train) even though the trains are more than one safe braking distance apart. Smooth operation, therefore, requires that fixed-block systems be operated with trains 2 or more blocks apart. This places a limit on the achievable headway with fixed-block headway control. Shorter headways are achieved by shortening the length of the blocks, i.e. by increasing the number of blocks per safe braking distance. The cost of one block (track circuit, receiver, transmitter, etc) is roughly $2000 up to a length of approximately 500 m. Consequently, to shorten the lengths of the blocks results in higher cost and lower reliability due to the increased equipment count.

5.1.1 Two-Block, 3-Aspect Indication Signalling

Two-block, 3-aspect indication signalling operates on a 2-block minimum train separation with 3 signal aspects: "Clear (G)", "Approach (Y)", and "Stop and Proceed (R)". The "Clear" signal permits a train to enter the block beyond the signal at the civil speed. An "Approach" aspect indicates that only one block in advance of the signal is unoccupied and that the train must be prepared to stop at the next signal. The "Stop and Proceed" aspect indicates that the block in advance of the signal is occupied and that the train must stop and then proceed (if safe) at a restricted speed. The length of the blocks in a 2-block, 3-aspect system is one safe braking distance or more. If only relatively long headways are run on the system, the block spacing can be increased to reduce the equipment cost. Figure 13 illustrates the operation of a 2-block, 3-aspect signalling system.

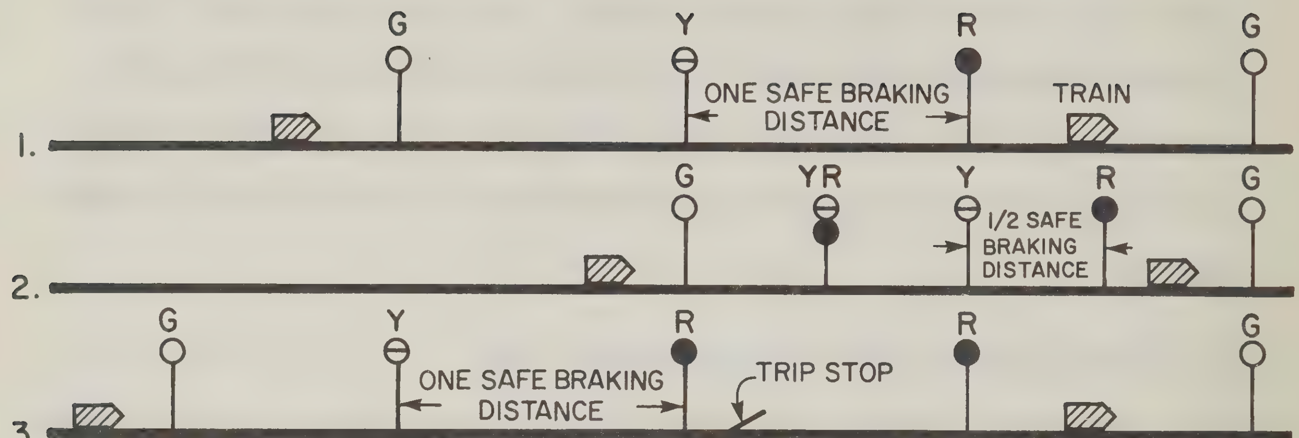


Figure 13; 1/ 2-Block Indication Signalling
2/ 3-Block Indication Signalling
3/ 2-Block Indication Signalling with 1-Block Overlap

5.1.2 Three-Block, 4-Aspect Indication Signalling

As the name implies, a 3-block, 4-aspect system operates with a minimum 3-block separation between consecutive trains. The blocks are usually made equal to 1/2 of a safe braking distance. Consequently, two trains following each other can be a minimum of 3 x 1/2 safe braking distance (i.e. 1 1/2 safe braking distances) apart.

The 4 aspects used are "Proceed (G)", "Advance Approach (Y)", "Approach (YR)", and "Stop and Proceed (R)". The "Proceed" aspect indicates that at least 3 blocks in advance are unoccupied. The "Advance Approach" aspect is only displayed if 2 blocks in advance of such signals are unoccupied. The "Approach" signal requires one unoccupied block in advance and the "Stop and Proceed" signal indicates that the block immediately in advance of the "Stop and Proceed" signal is occupied. Figure 13 shows the operation of this system.

5.1.3 Two-Block Indication Signalling with One-Block Overlap

This system differs from the 2-block, 3-aspect system in that the "Stop and Proceed" aspect is controlled through two blocks instead of one. This of course means that two consecutive trains may be no less than 3 blocks apart for safe operation without false alarm braking. The length of each block is equal to one safe braking distance*. The minimum possible train separation without false alarm braking is 3 x 1 safe braking distance. Figure 13 shows the operation of this system which is normally used on subways and other rapid transit systems operating at frequent intervals under manual control. This system works well with automatic trip stop devices since the extra block over which the "Stop" aspect is carried allows the trip stop to safely brake the

---

* One safe braking distance takes into account the worst case conditions multiplied by a safety factor normally taken as 1.35.

train before it could enter an occupied block.

5.2 Moving-Block Systems

The moving-block system is a headway (separation) control system which operates in the following way. When the separation of two consecutive vehicles is greater than 1 safe braking distance, the following train abides by the line velocity. When the separation between two trains is equal to 1 safe braking distance, the velocity of the following train is regulated so as to always maintain exactly 1 safe braking distance separation.

If one assumes that a failing vehicle stops instantly ('brick wall stop') the avoidance of collisions requires that the separation between vehicles should never be less than the maximum distance which a vehicle might require to execute an emergency stop (i.e., 1 safe braking distance). This then becomes the minimum headway achievable for any system operating under the 'brick wall' stop rule.

Whereas in fixed-block systems, position error may be as large as one block, moving-block systems can achieve relatively small position errors and can be operated very close to the minimum possible headway conditions. To utilize fully the capability of this system, automatic train operation is required.

5.3 Interlocking

Interlocking is a process which ensures that all signals and switches are interconnected in a manner which prevents unsafe train movements. This requires monitoring the position of all trains in the system, having knowledge of all switch positions, and of intended train routes. The information on the position of the trains is obtained directly from the train detection system. The switch status information is separately obtained but often use is made of coded track circuits to monitor the status of the track switches. With

present technology most, if not all interlockings are automatic and switches are remotely controlled.

5.4 Headway

As mentioned previously, "headway" in rapid transit systems usually refers to station through headway. Appendix A shows the analytical headway calculations under various headway control systems and for various vehicle characteristics.

6. AUTOMATION HARDWARE & IMPLEMENTATION

6.1 Automatic Train Protection (ATP)

The purpose of the Automatic Train Protection (ATP) system is to ensure safe separation between trains at all times. To accomplish this task the ATP performs the following functions:

1/ Surveillance of train and track

2/ Interlocking

3/ Overspeed protection

Figure 14 shows the functional relationship between the three ATP functions and train separation.

6.1.1 Train and Track Surveillance

6.1.1.1 Track Circuits

Track circuits are the most common means used for train and track surveillance. Track circuits are electrical circuits of which the rails of the track form a part. The operation of the track circuits is simple and is detailed in Section 6.1.1.1.2 (page 61). Track circuits are of two basic types: sections insulated from each other; and continuous track circuits which are used with continuous conducting rails. Figure 15 shows the further categorization possible within these two major types.

6.1.1.1.1 Insulated Track Circuits

Insulated track circuits consist of:

1/ Transmitter

2/ Receiver

3/ Rails

4/ Rail Isolators

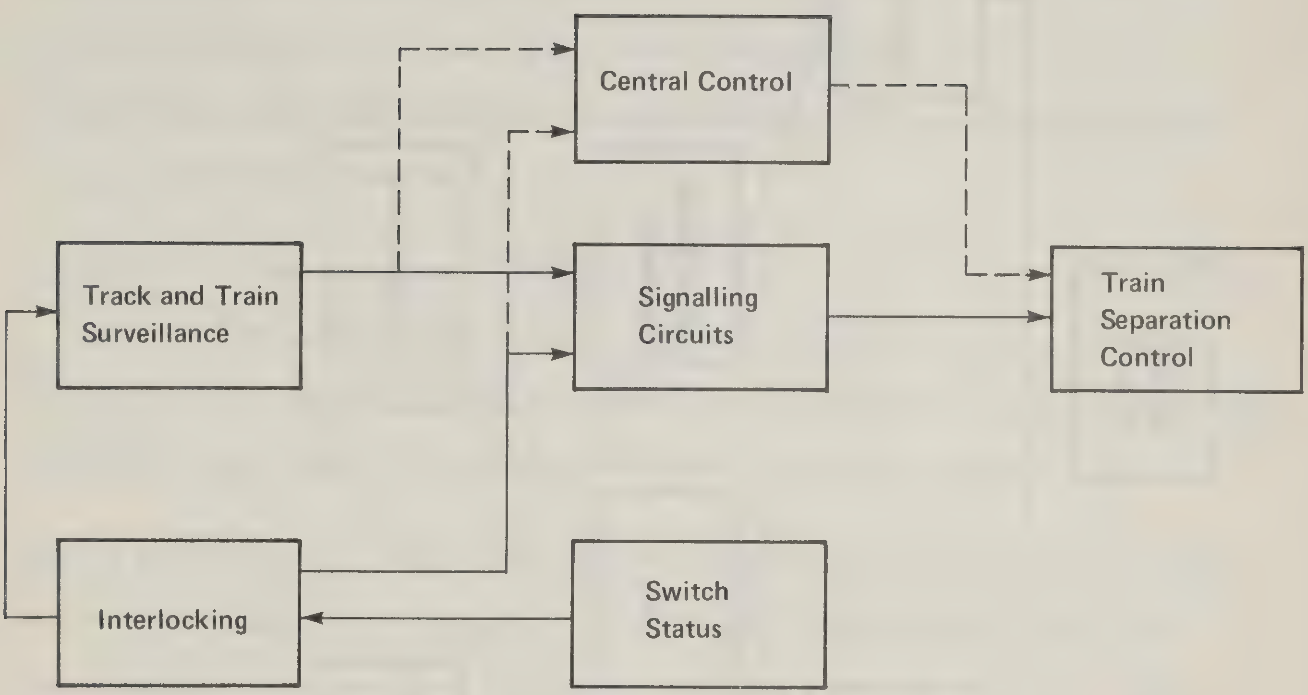


Figure 14; ATP Functions and Train Separation

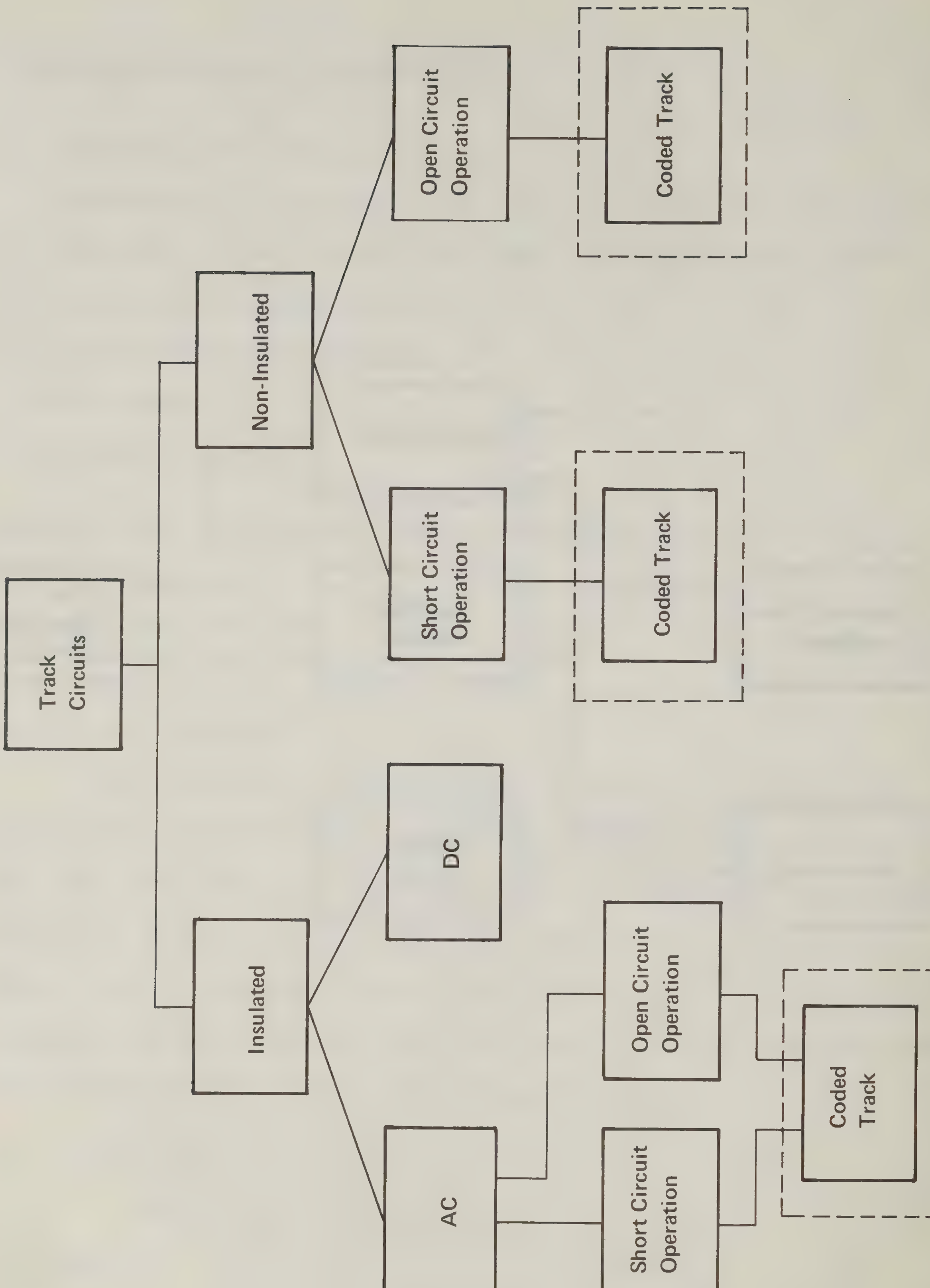


Figure 15; Track Circuits

The length of the track circuit is governed by the insulated joints built into the rails or, in the case of AC supply, by the rail impedance which increases with higher frequencies. Length limiting may also be enforced by frequency selective elements. Track circuits in rails laid on wooden or concrete sleepers are virtually short-circuited by vehicle axles entering the section so that the transmission between generator and receiver is interrupted. The interruption of transmission constitutes the criterion indicating the transition from a "line clear" to a "line occupied" condition. Since a broken rail results in transmission interruption, track circuits also serve a track surveillance function.

6.1.1.1.2 DC Insulated Track Circuits

DC insulated track circuits are continuously fed by low voltage (1.7 V to 2 V), well-regulated DC batteries which hold the track circuit relay (receiver) energized. Under this condition, the line is clear.

When a pair of wheels enters the isolated area through which the track circuit current is flowing, the current is shunted away from the track relay and the relay coils are de-energized indicating that the track is occupied. Since the circuit is normally closed, the failure of any component results in a restrictive condition; in other words, no failure can result in an unsafe condition.

Figure 16 shows the elements of a DC double rail track circuit. The advantage of DC track circuits is that they use very low power and that the track circuits are well defined by the insulation joints. The principal disadvantage of DC track circuits lies in the fact that the very low voltages are usually not sufficient to penetrate through the oxydization films that are formed on rails and that they depend quite heavily on the ballast impedance and the

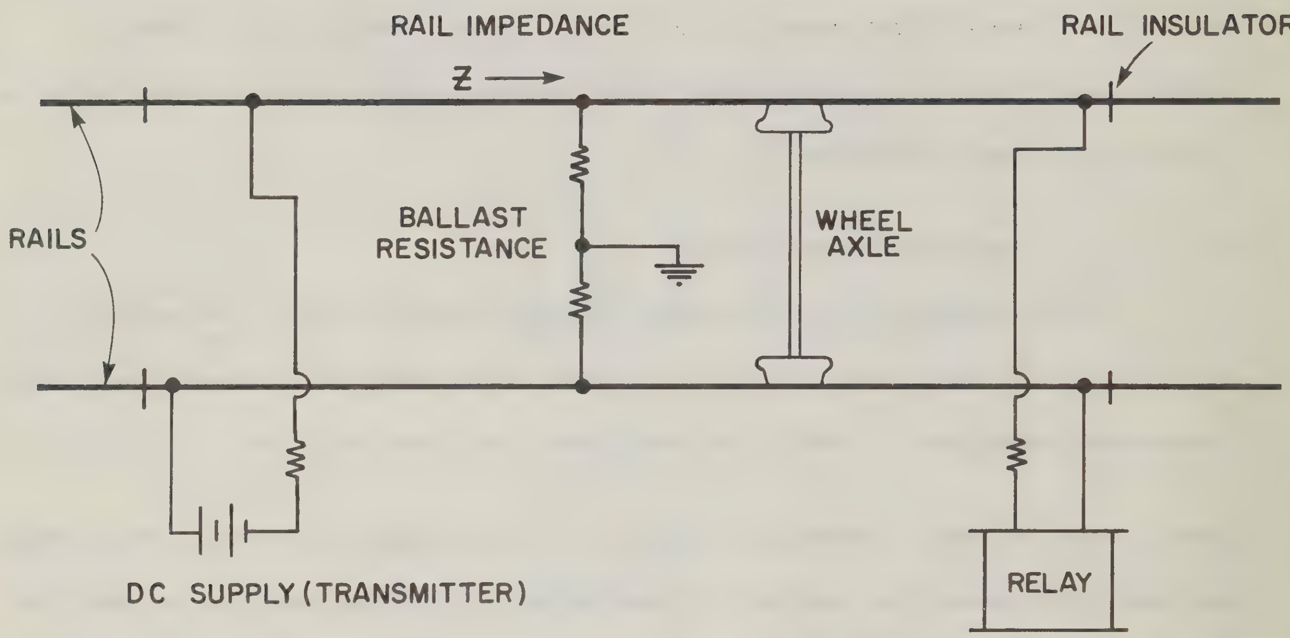


Figure 16; Insulated, Double Rail, DC Track Circuit

resistance of the axle. Particularly with light-weight vehicles, the shunting impedance of the axle-wheel-rail combination may not be satisfactory for reliable operations. Since no modulation is provided, the capability of DC track circuits is limited to a simple GO/NO-GO indication.

6.1.1.1.3 Insulated AC Track Circuits

Insulated AC track circuits operate under the same principles described for DC track circuits in the previous section. The main difference is that AC track circuits employ a constant-frequency transmitter instead of the DC battery.

Transmitter frequencies are 10 kHz or lower. The actual frequency chosen is a compromise between the rail attenuation, which increases with higher frequencies, and the wheel-to-rail capacitive impedance, which decreases with increasing frequencies.

If wheel-to-rail impedances or track oxidization are a problem, impulsive energy track circuits may be used. In impusive energy track circuits, the transmitter generates impulses having an amplitude of 20 V to 150 V and a duration of about 0.4 s at a repetition rate of 3 Hz to 8 Hz. Impulsive energy circuits should be satisfactory for light-weight vehicles since they can operate with much higher shunt sensitivities than simple AC or DC track circuits. The power consumption of these circuits is higher than for other track circuits.

6.1.1.1.4 "Open Circuit" Operation

In "open circuit" operation the relay (receiver) impedance is relatively large and this leads to a very low power consumption under normal conditions. The disadvantage of this approach is the very strong dependance on the ballast and axle impedance. As pointed out in the previous section,

impulse circuits and high supply frequencies minimize the problem of the axle impedance. The problem of the low ballast impedance (3 Ohm to 5 Ohm per 1000 ft (304.8 m)) and particularly the dependence of this paramater on weather does not permit a clear definition of a track circuit length which can be used under all conditions. (Typical ballast and rail impedances are shown in Table 9.) Though used extensively in the past, AC insulated track circuits operated in open circuit mode all suffer to various degrees from the problems described. Open circuit operation may be used with both insulated and non-insulated rails.

6.1.1.1.5 Short Circuit Mode

To avoid some of the difficulties associated with AC track circuits which are operated in the open circuit mode, the track circuit may be operated in a short circuit mode. Figure 17 shows a typical AC track circuit (insulated or non-insulated) which is operated in a short circuit mode. Operating in a short circuit mode minimizes the track circuit's dependence on the ballast impedance. Short circuit operation is achieved by tuning the receiver to the transmitter frequency. At the transmitter frequency, the impedance of the receiver circuit is practically zero (short circuit) and the ballast impedance in parallel with the receiver has no effect. This operational mode suffers from relatively high power consumption.

There are basically 4 coupling modes, i.e., modes for feeding energy for use with track circuits. Since these modes are similar for insulated and non-insulated AC track circuits, their discussion will be deferred to the next section.

Table 9; Typical Ballast Resistance and Rail Impedance

BALLAST RESISTANCE (TYPICAL)

CONDITION	BALLAST RESISTANCE
CLEAR BALLAST GOOD DRAINAGE	16.4 to 32.8 Ohm/km
FROZEN BALLAST	very large
TIES IN GOOD CONDITION NORMAL DRAINAGE	9.84 Ohm/km
CRACKED, DIRT FILLED TIES BALLAST IN CONTACT WITH RAIL	3.28 Ohm/km

RAIL IMPEDANCES (TYPICAL)

FREQUENCY	IMPEDANCE (Ohm/km)
25 Hz	0.035 + j 0.093
60 Hz	0.055 + j 0.19
100 Hz	0.067 + j 0.292
1 kHz	0.282 + j 2.68
5 kHz	1.22 + j 13.195

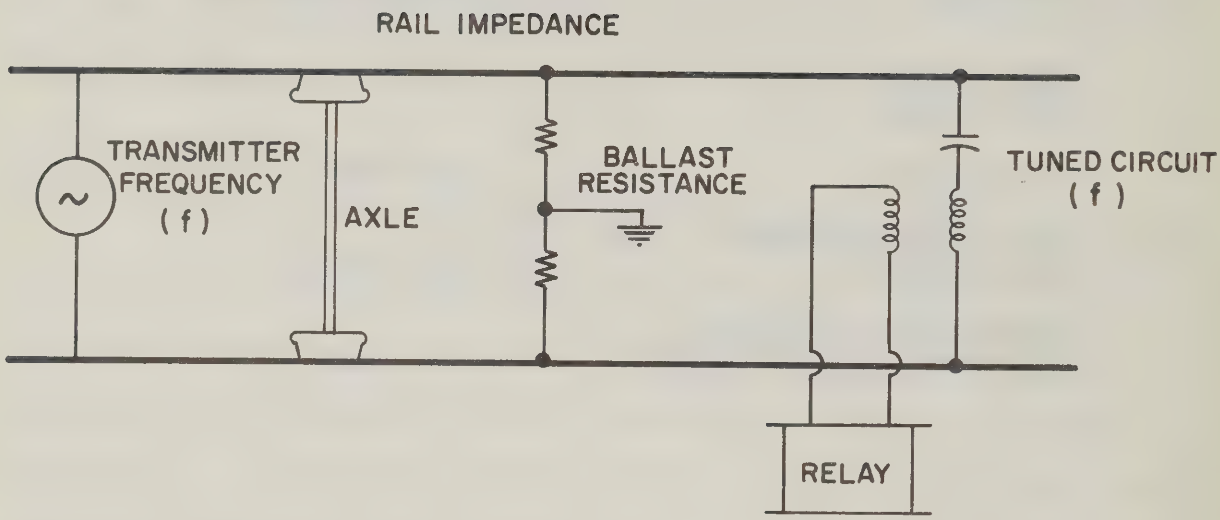


Figure 17; Non-Insulated, Double Rail, AC Track Circuit Operated in a "Short Circuit" Mode

6.1.1.1.6 Non-Insulated Track Circuits
(Jointless Track Circuits)

Non-insulated track circuits were introduced to be compatible with jointless rails. Since no physical insulation between rail sections is provided, adjacent track circuits must be electrically isolated from each other. Normally each section is terminated in an electrical short circuit (see also Section 6.1.1.5) and carrier frequencies are staggered from section to section. If tracks are closely spaced, different frequencies are used for the two tracks (see Figure 18).

The length of non-insulated track circuits is limited by the requirement for detection of broken rails and by a 0.06 Ohm shunt sensitivity*. It is clear that if the ballast were of infinite impedance, the current could not leak around the break; if the ballast were of zero impedance the track circuit could not operate. For impedance values between zero and infinity, and for a limited supply voltage, the ballast impedance defines the length of the track circuit within which broken rail detection and 0.06 Ohm shunt sensitivity can be provided. For short (50 m to 200 m) track circuits, sharp boundary definition is possible.

There are basically 4 combiniations of transmitter - receiver circuits used for non-insulated and insulated track circuits.

1/ Voltage Transmitted - Voltage Received

2/ Constant Current Transmitted - Voltage Received

3/ Constant Current Transmitted - Current Received

4/ Constant Voltage Transmitted - Current Received

---

* For light rail vehicles much higher shunt sensitivities of about 1 Ohm are required.

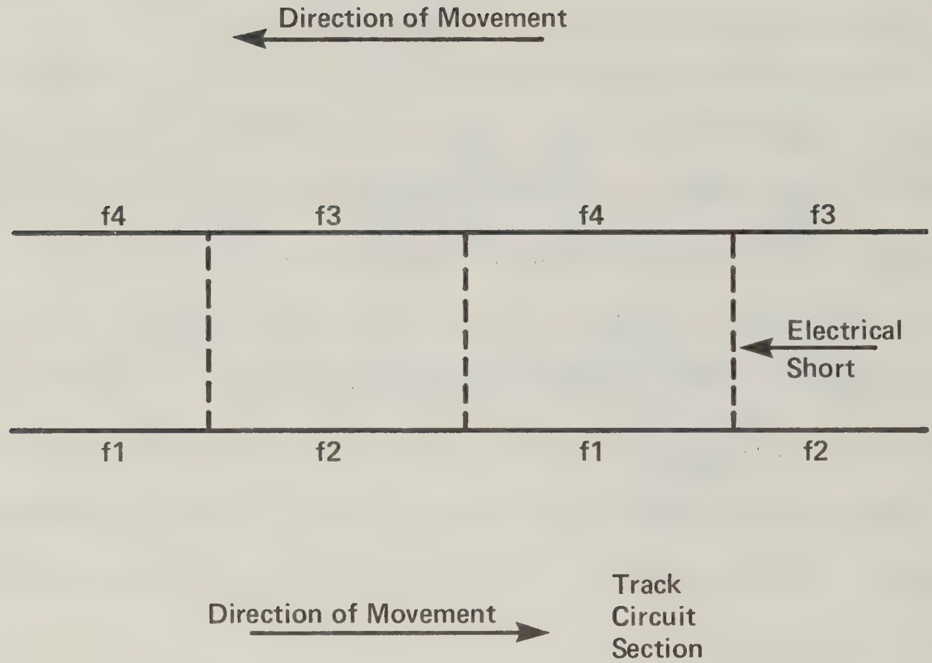


Figure 18; Staggering of Track Circuit Frequencies
(f1, f2, f3, f4, one track circuit frequencies)

Each of these combinations has different performance characteristics, the most important of which are shown in Table 10.

6.1.1.1.7 Coded Track Circuits

The AC track circuits previously described operated on a single transmitter frequency and could only indicate line clear or line occupied status (broken rail status is also included), however, additional information may be transmitted via track circuits by modulating the carrier frequency. For instance, the AC-fed track circuit may be modulated at 180 Hz/min for a 'line clear' indication, at 120 Hz/min for an 'approach medium' indication, at 75 Hz/min for an 'approach' indication, and at no current for a stop aspect.

Normally no more than 10 different pieces of information are transmitted by coded track circuits. To transmit more information requires the use of higher frequencies which are not compatible with the attenuation characteristics of the rail. Coded track circuits can be amplitude or frequency modulated. Also, "2 of 5" codes have been used successfully. These codes allow 10 different messages and yield greater information security since for any message to be recognized as valid, 2 of 5 distinct frequencies must be received and decoded. Additional information transmission need not be restricted to train commands but can be used to set signals and interlocks.

6.1.1.2 Axle Counters

Axle counters are used for train surveillance. The operation of these devices is relatively simple. At the beginning and end of each track section an impulse counter is buried in the track. As the train passes over the track section each of the train axles generates an impulse which is counted and memorized by the axle counter. As long as the number of axles which have entered a specific track section exceeds the number of axles counted at the

Table 10; Track Circuit Characteristics vs. Transmitter and Receiver Characteristics

Transmitter	Receiver	Boundary Definition	Coupling Losses	Length of Track Circuit
Voltage	Voltage	50 m	low	1000 m
Constant Current	Voltage	50 m	medium	1000 m
Constant Current	Current	5 m	high	500 m
Voltage	Current	10-20 m	medium	1000 m

end of the section, the track section is considered occupied. To satisfy operational requirements the axle counters must operate down to zero speed and must be able to recognize reverse direction operation.

Axle counters do not provide broken rail indication and provide a maintenance problem since they are buried in the track. Another significant shortcoming of these devices is that they do not provide continuous detection of the vehicles. A train or vehicle is only counted (detected) at discrete points and is assumed to be within that section until an equal number of axles have been detected to emerge from that section. Axle counters are based on mechanical, inductive, and electronic principles.

6.1.1.3 Inductive Impedance Circuits

Inductive impedance circuits are used to detect the presence or absence of trains along the track. The system uses wire loops installed between the rails. The magnitude of the loop impedance depends on coupling - the presence or absence of a train above the loop. The loops which divide the track into track sections are connected to an impedance bridge and scanned cyclically.

The advantages of this system are that it is independent of the unreliable wheel/rail interface and that it offers continuous train detection. However, broken rail protection is not provided and the circuit can only transmit train presence or train absence information.

6.1.1.4 Continuous Conductor Type Train Surveillance

Due to the attenuation effect of the rail, track circuits are limited to low frequency operations. Low frequencies, however, also mean low data rates. To increase the data rates, alternative methods of information transmission have been developed. One of these is the continuous conductor

type system which has been developed in Great Britain and Germany. It consists of insulated conductors placed between the rails and transposed at regular intervals. A feeder (transmitter) and a receiver are directly connected to the conductors. The feeder transmits frequency shift (FSK) modulated information to the trains at a carrier frequency usually between 20 kHz and 100 kHz and the trains, which are equipped with receiving and transmitting antennae, first receive the information transmitted and then reply at a different carrier frequency. The reply messages are coupled into the conductors and received by the wayside receiver. The conductor crossings are used to determine the position of the train relative to a "landmark".

A large amount of information can be transmitted in this way. Data rates are usually 1200 bits/s or lower, although higher data rates (2400 bits/s) have been considered. Train surveillance is closed loop since the train detects loop transpositions and transmits this position information via the conductors to the central control where it is evaluated with respect to a preprogrammed profile. Absence of information or anomalous information results in a shutdown.

The advantages of continuous conductor train surveillance systems are obvious. Firstly, they combine train detection and information transmission into one system. This permits much higher data rates than would be possible by track circuits. They offer independence from the troublesome wheel/rail interface and maintenance of conductors can be simple since they are not buried in the track (however, accidental damage to conductors may occur frequently). If the loops (i.e., cross-overs) are frequent, the detection method can be considered continuous.

The shortcomings of this system are:

1/ No broken rail protection

2/ No absolute position indication

3/ No independence from train-mounted equipment

6.1.1.5 Summary

Section 6.1 has presented a brief review of the various train surveillance systems currently available. Basic system operations were explained and illustrated. Advantages and shortcomings are highlighted and summarized in Table 11.

6.2 Automatic Train Separation Control

Figure 19 shows a very simple division of automatic train separation systems. "Automatic train stop" refers to systems which in case of a stop signal violation, automatically activate the train brakes. "Cab signalling" refers to any system which displays a signal in the driver's compartment indicating a condition affecting the movement of the train and advises the driver prior to applying the brakes. Whereas automatic train stop is a means of enforcing obedience to a stop signal, cab signalling permits better all-weather operation of the trains and can be used to independently enforce obedience to signals.

6.2.1 Automatic Train Stop Systems

6.2.1.1 Mechanical Automatic Train Stop

Mechanical automatic train stops are located along the wayside. There is a connection between wayside signals and the automatic trip-stop. If the wayside signals display a red or stop aspect, the mechanical arm of the train stop is raised. Should a train attempt to run through a signal which displays

Table 11; Summary of Train Surveillance System Characteristics

Parameter / Surveillance System	Information Bandwidth	Frequency	Detection	Broken Rail Protect.	Wheel/ Rail Depend.	Applications
Track Circuit, DC	very low	DC	continuous	yes	yes	heavy, non-electrical trains, low information rate
Track Circuit, AC	low	to 10 kHz	continuous	yes	yes	heavy trains, low to medium information rate
Impulsive Energy Cct.	low	8 Hz	continuous	yes	not very signif- icant	low to medium rates, light rail vehicles, relatively short distances (energy consumption)
Axle Counters	very low	N/A	discontin- uous	no	not very signif- icant	low data rates, long headways
Loop Impedance	very low		continuous	no	no	where independence from wheel/rail interface is required, low information rate
Continuous Conductor	high	20 kHz - 100 kHz	continuous	no	no	high data rates, indepen- dence of wheel/rail inter- face, short headways

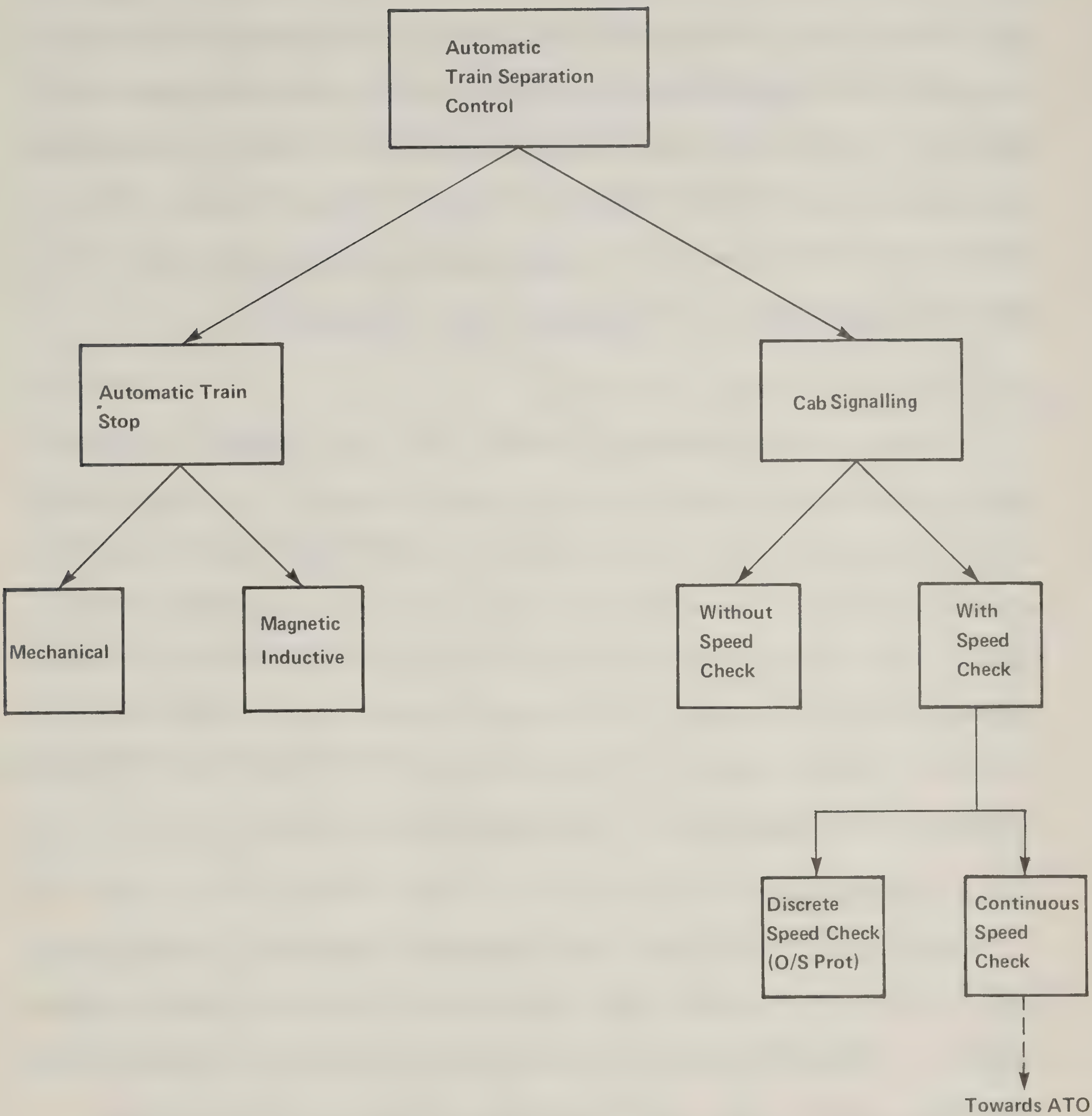


Figure 19; Automatic Train Separation Systems

a stop aspect, the train stop's mechanical arm trips the brake valve on the train and the brakes are automatically applied. These devices are rugged and simple but not well suited to high speed systems and inclement weather use. For short headway applications the actuation time of the device must also be considered. There is no speed checking feature on this type of equipment and most systems permit low speed "keying" through. This key-through feature permits deliberate misuse of the system, but at the same time it is useful in permitting operations in case of signal or trip-stop failure.

6.2.1.2 Magnetic Automatic Train Stop

Similar to the mechanical automatic train stop, magnetic automatic train stop devices are connected to the wayside signals. If a train passes over an active track magnet in violation of a stop signal aspect (magnets are normally activated if the associated signal displays a stop aspect), the magnetic force field in the train reverses the main relay contact and the emergency brake is activated. The enforced braking can be cancelled by a release key (key-by feature). Though better suited to high speed applications, the buried track magnets always pose a maintenance problem.

6.2.2 Cab Signalling

Cab signalling is a system that provides an indication of the wayside signal aspect to the driver right inside his cab. Normally, "speed checking" refers to a system in which an indication inside the cab tells the driver the permissible speed. Often, speed checking is combined with an overspeed protection system which automatically applies the brakes should the train exceed the permissible speed. In the following sections, "speed checking" shall refer to the provision of a permissible speed indicator inside the cab combined with an overspeed protection system.

Cab signalling with speed check is a first step toward fully automated operation. Cab signalling may have continuous or discrete speed check capabilities as shown in Figure 19. Whereas most systems apply the emergency brake in the case of overspeed condition, some systems use the service brake, releasing it when the permissible speed has been reached. For this approach to be safe, the service braking system has to be fail-safe.

6.2.2.1 Cab Signalling Without Speed Check

In its most elementary form a cab signalling system consists of the following:

1/ Wayside Transmitter

2/ Pick-up Coil

3/ Vehicle-Cab Signal Receiver

4/ Control Device

5/ Alarm Bell and Display

6/ Reset Button

7/ Equipment Status Indicator

The wayside transmitter may work on magnetic or electro-magnetic principles. (The Indusi type equipment used extensively in Germany is an example of the former and the "S" Type Cab Warning Device of the Japanese Railways is an example of the latter.) Whichever principle is used, the basic block diagram of an intermittent cab signalling device takes on the form shown in Figure 20.

The device operates in the following manner. The status of a wayside signal commands the information to be transmitted by the cab signal transmitter. The cab signal transmitter signal is received by the vehicle pickup coil then amplified, filtered, and decoded in the decoder unit. When a vehicle passes over a transmitter which repeats a "clear" aspect, the cab unit is kept in the

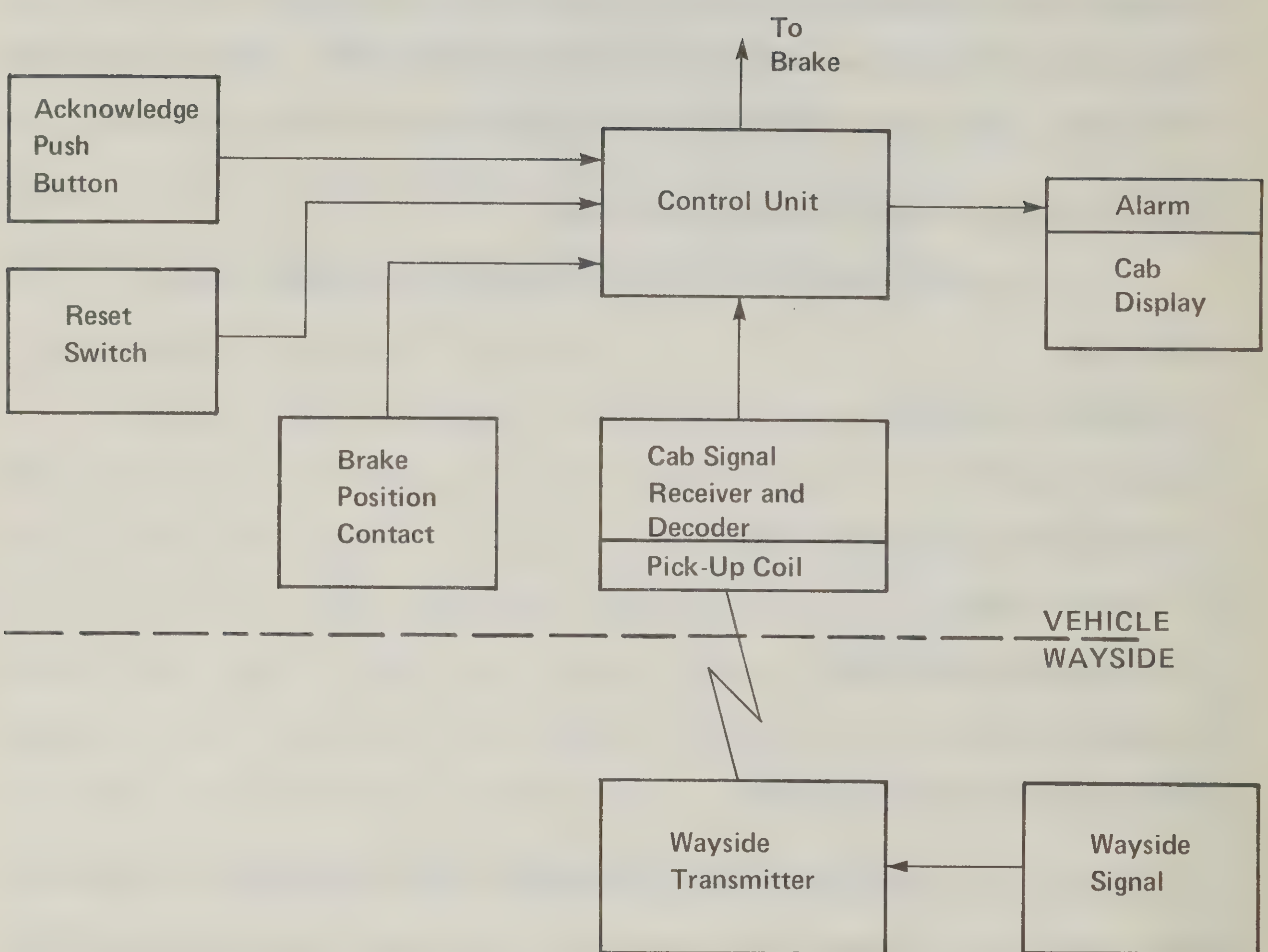


Figure 20; Cab Signalling Equipment

normal condition. When a vehicle passes over a transmitter which repeats a "stop" aspect, the decoder instructs the control unit which, in turn, energizes an acoustical and a visual warning device. The driver is given 3 s to 6 s in which to start braking and to acknowledge the alarm condition by pressing the "acknowledge" button - otherwise, the emergency brake is automatically activated.

For intermittent cab signalling to be effective the cab signal transmitter must be placed a distance, S, upstream of the wayside signal where "S" is given by:

$$S = (5\ s^{*} \times V_{max}) + \text{braking distance} + \text{safety margin}$$

This requirement, however, restricts the use of intermittent cab signalling without speed check to applications with great intersignal distances and no requirement for speed checking and/or enforcement.

The advantages of these train protection systems are:

1/ High reliability with low equipment complexity

2/ Low cost

3/ Easy expansion to include speed checking and control

These advantages are counterbalanced by:

1/ Useful only where long (3 min and up) headways are required

2/ Maintenance problems due to wayside buried equipment

3/ Reduced operational efficiency of the system

---

* Assuming the acknowledgment time limit for the driver is 5 secs.

6.2.2.2 Cab Signalling with Speed Checks

Fundamentally, an automatic train separation control system must accomplish its function within the following two requirements:

1/ It must enforce a safe separation between all vehicles at all times

2/ It must not adversely affect the operational efficiency of the transit system

Automatic train stops and discrete (intermittent) cab signalling systems without speed checking generally meet the first requirement, however, they fail to meet the requirement of unhindered operational efficiency. Speed checking and automatic braking in overspeed situations allows higher operational efficiency since obedience to the signals is enforced at all times and the probability of high speed emergency stops is avoided. In the following paragraphs, two methods of speed checking, one discrete and one continuous, will be described and compared.

6.2.2.2.1 Cab Signalling, with Discrete Speed Checks (Overspeed Protection)

Though there are a number of different implementations (see Table 12), basically all systems of this type rely on comparing the train speed with the permissible speed at certain spots along the track. These spots are strategically located to permit the vehicle to reduce the speed or come to a halt according to the signal aspect displayed.

A train moving along the track undergoes spot speed checks and if it passes over any check point at a speed exceeding a specified limit, the brakes are automatically applied. Since the speed is checked at several levels, the automatic release of the brake when the train has slowed to a certain speed compiles with safety requirements.

Table 12; Cab Signalling With Discrete Speed Check

Features / Type	Signal Transmission Wayside to Train	Frequencies	Method of Speed Check	Special Features	Applications
Two Coil Single Frequency	Resonant Coils in track are activated by restrictive wayside signal and are detected by on-board receiver	100 - 130 kHz	Time to travel between spaced ground coils. Measures average speed.	.easily installed .no speed limit value transmitted to the train .since average speed is measured, system is not safe against bad driving	.relatively few aspects .shorter block .can be used as speed check
Single Coil, Multiple Frequency	Resonant Coil is set to resonate at different frequencies according to signal aspect indications.	60 - 80 kHz 100 - 140 kHz	Comparison between speed limit represented by resonant coil frequency and running speed (tachometer) Measures Instantaneous Speed	.zero speed check possible .instantaneous speed checking requires judiicial choice of number and placement of speed checks to be safe.	.large block distances

Figure 21 illustrates the operation of a typical discrete cab signalling system with discrete speed checks. As can be seen, the speed is checked at certain spots only, with the train running without speed check between the speed check points. Since short block separations (i.e., short headways) imply strict speed limit enforcement, cab signalling with discrete speed checks is not well suited to these applications.

The system described above attempts to measure the instantaneous speed of the vehicle as it passes over the speed check spots, however, there are other systems which measure the average speed of the vehicle.

The 2-coil system (Table 12) and the system described below are examples of the average speed method. The simplified schematic diagram shown in Figure 22 shows an inductor type speed control system aimed at checking the average vehcile speed.

The location of the inductors buried in the track is important. The P1 inductor is placed at the start of the controlled section; the BX inductor is placed at the end of the controlled section; and the PIF resetting inductor is downstream of the BX inductor. With no train in the control section, the BX inductor is tuned to the stop frequency.

When a vehicle passes over the P1 inductor at the start of the control section, the signal from the inductor triggers a fail-safe memory circuit. The memory feeds a timer. After the lapse of the time T, the timer readjusts the frequency of the BX inductor. If the speed of the train is too high, it will reach the BX inductor before its frequency has been readjusted. The train is then automatically brought to a stop. When the train reaches the PIF reset inductor, the switching signal for the BX inductor disappears and it is tuned to the stop frequency again.

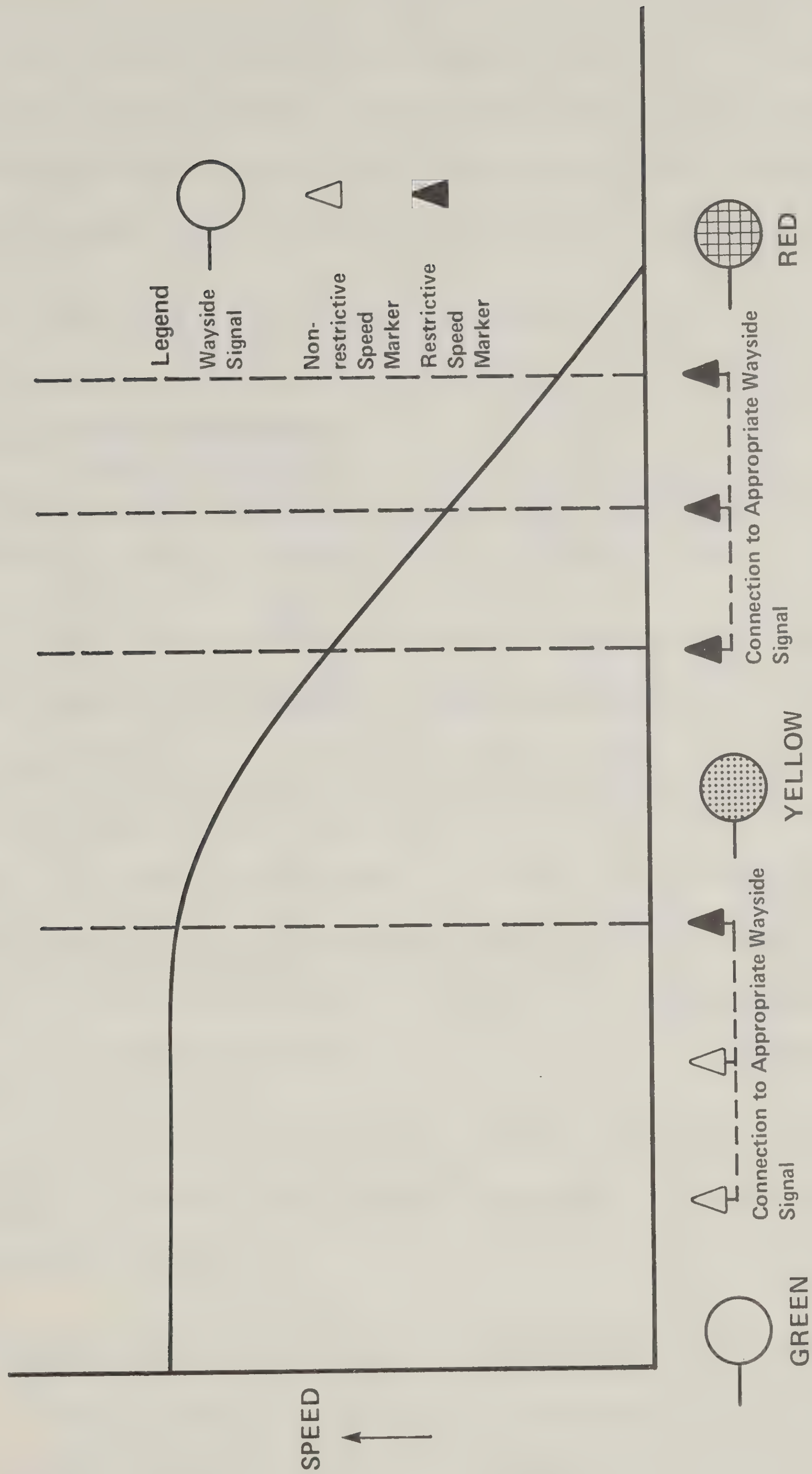


Figure 21; Typical Cab Signalling Process with Discrete Speed Check

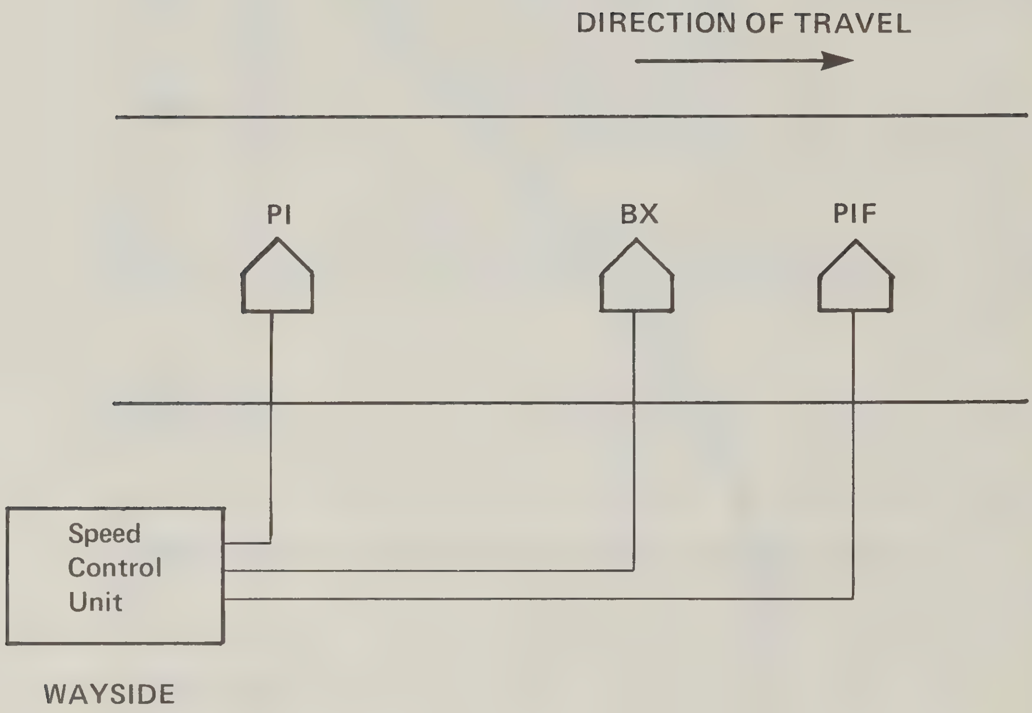


Figure 22; Speed (Average) Checking System Schematic

A general block diagram of the on-board equipment is shown in Figure 23. The operation of the circuit has been previously described in this section under discrete cab signalling. In addition to checking the vehicle speed, the system also protects against overspeeding at discrete points along the track.

6.2.2.2.2 Cab Signalling with Continuous Speed Check

This system supplies the driver with the following information:

1/ The distance between the train and the most restrictive signal

2/ The permissible speed of the train at the next restrictive obstacle

3/ The permissible speed of the train at each instant in time

The driver is provided with advance information (speed at next obstacle) which allows him to operate the train even at high speeds. Since in the case of continuous speed checking, errors cannot accummulate over the sampling period (as is the case with discrete speed checking), tolerances on train separations can be closer and shorter headways can be achieved. The amount of information required for cab signalling with continuous speed checking begins to exceed the bandwidth of track circuits and many systems make use of inductive loops which have a much higher bandwidth.

To enhance the reliability of the system, wayside transmitters (and receivers) are duplicated and provided with automatic changeover in case of failure. On-board equipment is triplicated.

A typical display indicator as used in cab signalling with continuous speed checking is shown in Figure 24. The driver is advised at all times as to his permissible speed, the actual speed of his vehicle, the distance to the next obstacle (speed restriction), and the speed permitted at the next obstacle. In contrast, discrete speed checking only compares the vehicle speed and the

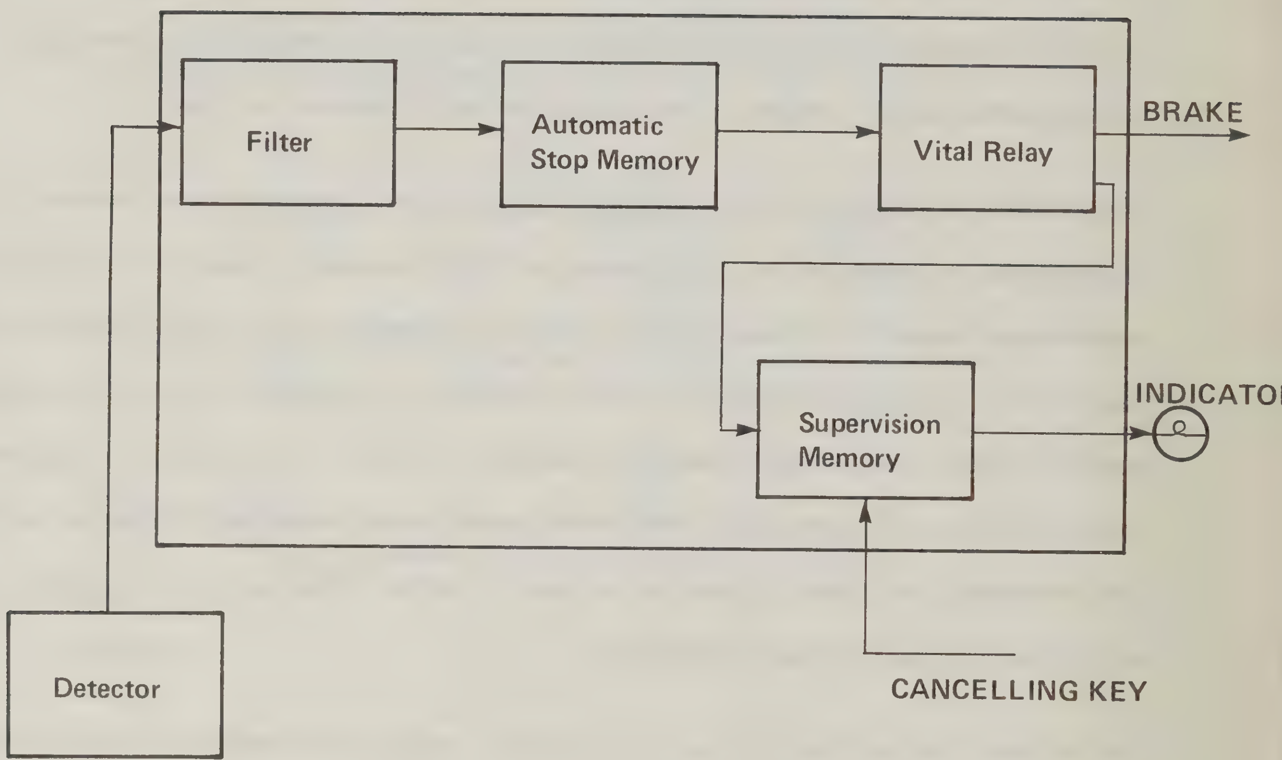


Figure 23; On-Board Equipment – Discrete Speed Check and Overspeed Protection

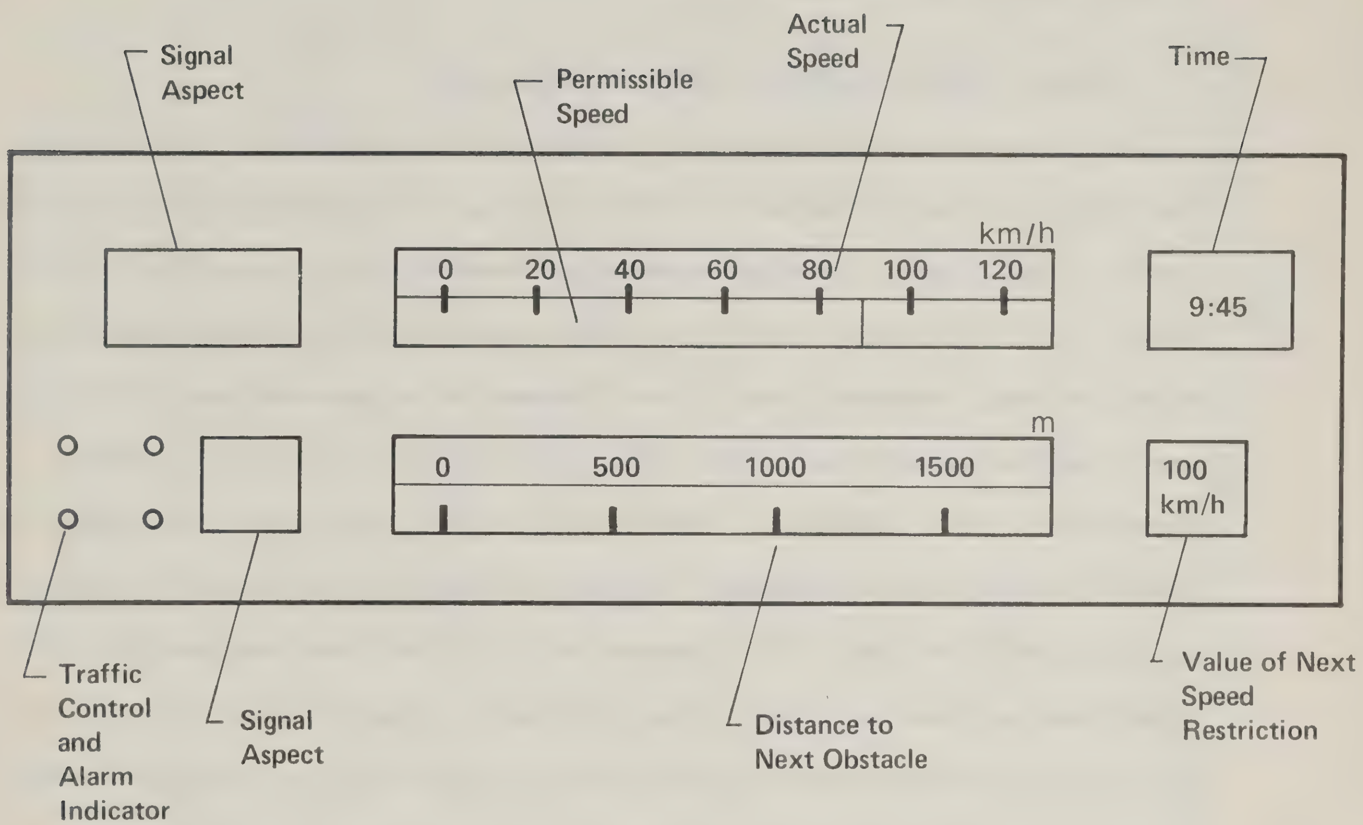


Figure 24; Display for Cab Signalling with Continuous Speed Checking

permissible speed at various check points.

6.2.2.2.3 Summary

Table 13 briefly summarizes the main characteristics of Automatic Train Separation protection systems.

6.3 Automatic Train Operation (ATO) Architecture

The functional requirements of Automatic Train Operation (ATO) systems have been previously discussed. To reiterate, the fundamental task of the ATO system is to function as the automatic motorman which controls the tractive effort. It will become clear that a cab signalling system with continuous speed checking possesses some of the fundamental elements of an ATO. The step from cab signalling to ATO, which technically is relatively simple, represents a very fundamental departure from past practice. It replaces the human operator with an automatic system and relegates the driver to monitoring and vigilance functions. ATO systems can be divided into narrow-band and broad-band systems according to the information channel bandwidth that they use. This division could also have been introduced at the ATP level, since the division arises from a channel capacity consideration, but ATO represents the first level at which broad-band and narrow-band channel capacities lead to significant differences in the train control system architecture.

6.3.1 Narrow-Band ATO

Narrow-band ATO systems all make use of track circuits to transmit messages and commands from the wayside to the train and they are, therefore, restricted to the narrow bandwidth offered by the rails. As noted previously, track circuit frequencies are usually in the lower audio frequency range (below 10 kHz) and they are usually amplitude modulated at frequencies below 22 Hz. Higher modulation frequencies are used for high-speed or short-headway

Table 13; Characterization of Automatic Train Separation Protection Systems

Parameters Type of Train Separation Protection	Environmental Dependency	System Speed (m/s)	Headway (s)	Potential for Conversion to ATO
<u>Automatic Train Stop</u> Mechanical	Medium to High	Medium (25 m/s)	90 + s	low
Magnetic/Inductive	Low	30-35 m/s	90 + s	low
<u>Cab Signalling</u> Without Speed Check	Low	30-35 m/s	90 +	Medium
With intermittent Speed Checks	Low	to 50 m/s	60 +	Medium
With Continuous Speed Check	Low	to Max. Achievable	45* + s	High

* depending on vehicle length and vehicle performance

trains which require more signal aspects for their operation. To enhance safety, codes using 2 out of 5 allowing 10 commands, and 3 out of 7 frequency combinations allowing 35 commands are transmitted to the trains. Only if valid frequency combinations are received is the message accepted by the trains.

The ATO equipment is divided into wayside and on-board equipment, each of which will be discussed in the following section.

6.3.1.1 On-Board Narrow-Band ATO Equipment

Figure 25 shows a simplified schematic of a narrow-band ATO on-board equipment. The heart of the on-board ATO equipment is the regulator. The function of the regulator is to control the traction effort in accordance with the command signals and within the required jerk and acceleration envelope. When running between stations, the regulator receives its commands from the Automatic Train Protection (ATP) system which is usually based on a fixed-block system, with AC track circuits and on-board decoders. Near the stations, the input to the regulator is transferred from the ATP system to the Stopping Program Reference Generator System which generates an accurate stopping profile for submission to the regulator. The ATP, however, retains an overriding capability to prevent rear-end collisions.

Between stations, the regulator takes the ATP permissible speed command, compares it to the actual train speed (as measured by a tachometer or other velocity measuring device) and produces a signal which consists of the algebraic difference between the commanded speed and the actual speed. This difference signal is amplified and applied to the power conditioning unit. Train acceleration and jerk are automatically limited to accepted passenger comfort criteria.

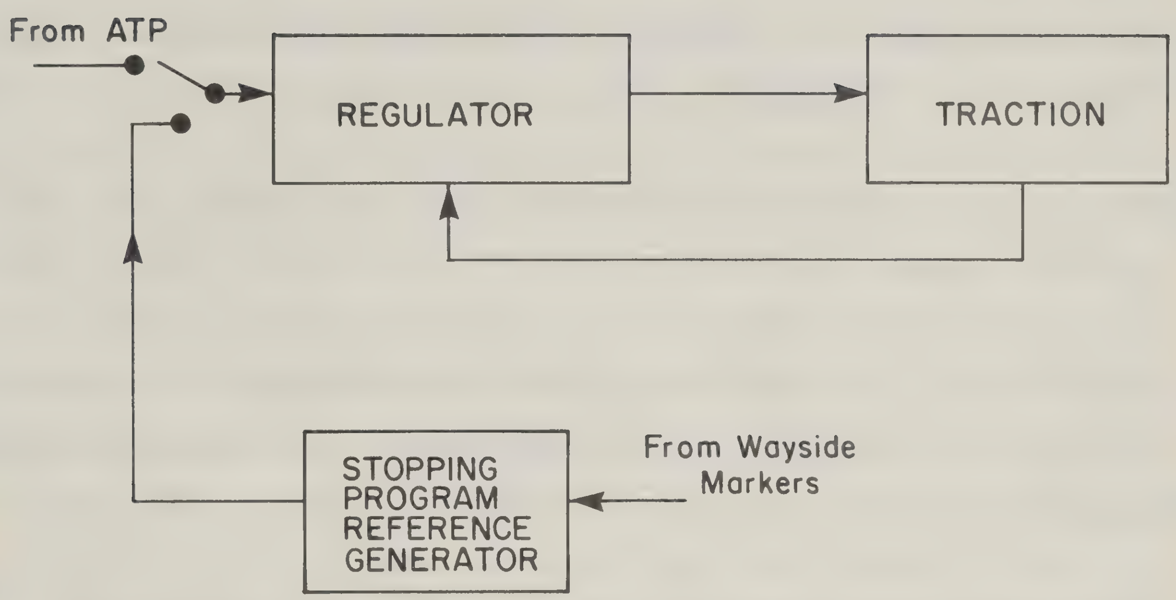


Figure 25; Narrow Band ATO

It is noteworthy that narrow-band ATO systems employ a "stand-alone" type of automation, i.e., they clearly separate each of the functions to be automated and implement them individually. The advantage of stand-alone automation is that failure of one particular subsystem such as the ATO or ATS does not affect the normal operation of the remaining subsystems. Strictly speaking, this is true only if either ATS or ATO has failed, since a failure of the ATP would not permit system operations. It is for this reason that ATP systems are built with very high reliability. A mean time between failures (MTBF) of 10,000 h is a reasonable design goal for ATP systems*.

The reliability required of the narrow-band ATO system depends to a large extent on whether a driver is retained on board the vehicle or not. With a driver on board, a lower ATO reliability can be accepted since in many failure situations the driver can run the train manually under ATP supervision. (This question of ATO reliability required will be raised again in the economic analysis of automation.)

Not all narrow-band systems are designed to perform all of the possible ATO functions (these are listed in Chapter 3). In fact, there is such variety in the number of functions actually carried out by the ATO that a categorization is not possible. Some systems carry out the door opening and closing function through the ATO and the only driver function is to press a start button for train departure. Other systems require the driver to perform the door closing and opening functions. Still other systems run under ATO between stations but relinquish control to the driver in the vicinity of the stations. In any event, the on-board ATO equipment for all narrow-band ATO systems is quite similar and

---

* This figure, however, is almost never achieved in practice and actual figures are in the 1000 h range.

only minor functional changes occur.

The on-board ATO costs also vary from system to system. Typically, they run $5,000 to $15,000/vehicle (1974 dollars)* if lower reliability (1000 h MTBF) is acceptable, and $25,000 to $30,000/vehicle if high ATO reliability is required.

6.3.1.2 Wayside Equipment (Narrow-Band ATO)

Wayside equipment requirement for narrow-band ATO is not very extensive. Figure 26 shows a possible schematic of narrow-band ATO wayside equipment which consists mainly of markers (signal spots) for the Stopping Program Generator. The markers are located at the wayside near the stations and consist of tone transmitters which provide a reference input to the Stopping Program Generator. Normally, 2 or 3 markers are used, however, some systems require more.

It must be realized that most of the wayside costs are ATP wayside equipment costs (track circuits, etc.) and that ATO wayside equipment costs are usually not very large. A typical cost figure for total ATP plus ATO wayside equipment is $300,000/km of single track. Of this figure, approximately 10% can be directly attributed to ATO, which leaves $270,000/km (single track) for ATP and $30,000/km (single track) for narrow-band ATO wayside equipment.

In addition to the stopping program markers, station/vehicle door interface equipment is required to provide absolute berthing information to the vehicle. If driverless operation is foreseen, the station/vehicle door interface must also regulate the passenger boarding and disembarking safely and efficiently.

---

* On-board ATO costs for the Montréal Métro were roughly $15,000/vehicle with an MTBF of 500 h.

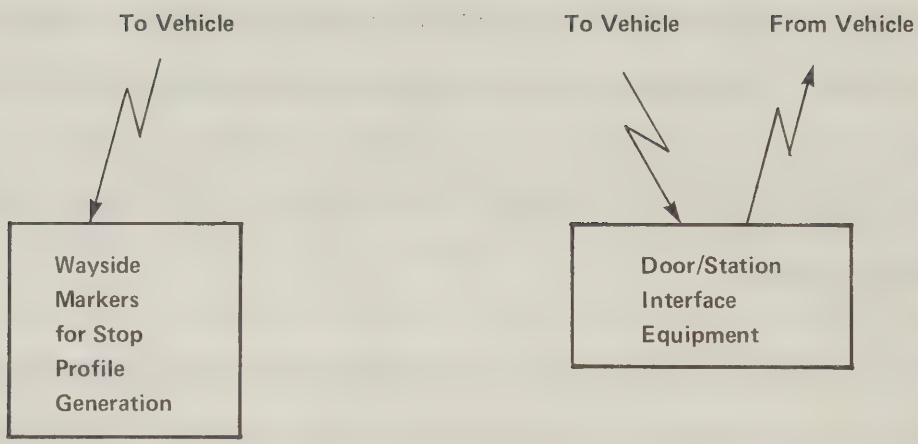


Figure 26; ATO Wayside Equipment

6.3.2 Broad-Band ATO

The term "broad-band" refers to the frequency bandwidth of the system. Broad-band systems have higher information capacities than narrow-band systems which are characterized by "stand alone" type of automation where each subsystem performs its own function separately. Broad-band systems, in contrast, are characterized by a high level of integration of the various functions.

Among the advantages of broad-band systems is the ability to exchange large amounts of information and therefore effect a finer control and the potential of more compact system design. The greatest disadvantage of this type of system is that it requires fail-safe or checked redundancy techniques throughout the system since circuits are "shared" by ATP, ATO and ATS. Also, design techniques for broad-band systems are less proven and formalized than for narrow-band systems.

6.3.2.1 ATO Broad-Band Equipment

Figure 27 shows the ATO and ATP intermingling. This is typical of broad-band systems. The figure shows both wayside and on-board equipment. The "running and brake controller", which is located on the wayside, receives all track occupancy information (ATP function), all signal and route interlockings, switch status messages, and characteristics of the route. Based on all these inputs and on the headway strategy, the controller produces the necessary ATO commands. The commands are always constrained by safety (ATP) considerations. The wayside transmitter transmits the ATO messages to the train and the receiving antenna on board the train captures the message. The signal from the vehicle receiving antenna is fed through amplifiers and filters to a decoder; the decoded message is checked (ATO and ATP functions), and if all checks prove satisfactory, an authority to proceed signal is generated.

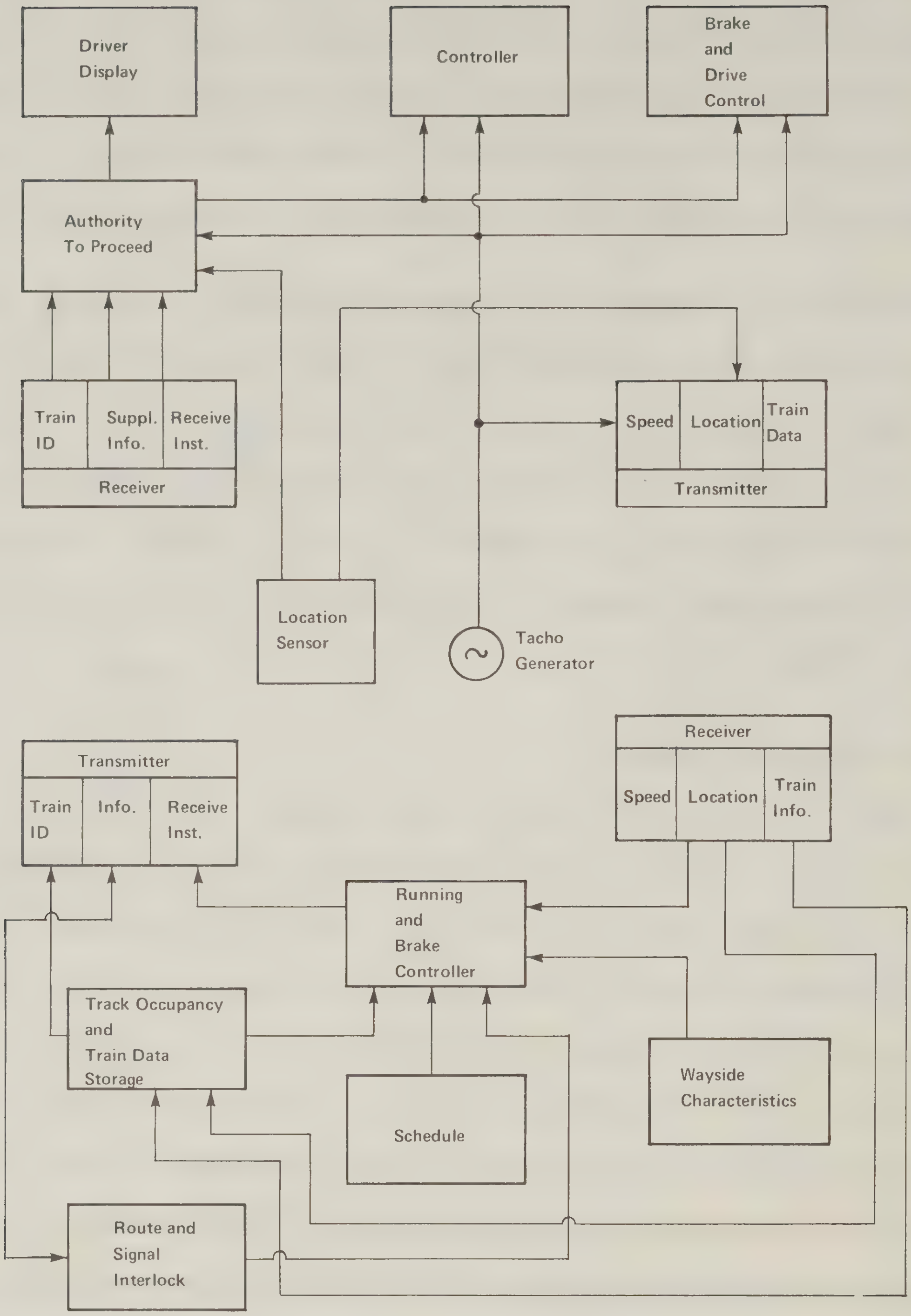


Figure 27; Block Schematic, Broad Band Train Control

The wayside-to-train messages comprise an address to locate the train, acceleration, coasting, and braking commands, the maximum permissible speed, braking distance, target speed and other supplementary information. The wayside transmits at one frequency and the trains reply at another. The modulation is by frequency shift keying (FSK) and the data rate is in the order of 600 to 1200 bits/s.

For systems which retain a driver on board, manual operation is facilitated by supplying the driver with the following information:

1/ permissible speed

2/ speed "restriction" of the next obstacle

3/ distance to the next obstacle (train, signal, etc.)

This amount of information continuously supplied exceeds the capacity of narrow-band systems. In broad-band systems, manual operation can readily maintain service level with only minor degradation under adverse weather or failed ATO conditions because sufficient information is supplied to the driver to operate his train.

Under automatic train operation, the train responds to wayside messages with return messages of its own. These return or train-generated messages contain the position of the train, the speed of the train, percentage of braking effort, train length and other supplementary information. It should be noted that in broad-band ATO systems the control loop which controls the longitudinal motion of the vehicle is closed through the wayside equipment. This is in sharp contrast to the narrow-band ATO system which closes the train longitudinal control loop on board the train. As a result, narrow-band systems do not directly supply train speed to the wayside and cannot achieve quite the same control accuracy as broad-band systems which continuously report to wayside

position and speed.

6.4 Automatic Train Supervision (ATS)

The functions of the Automatic Train Supervision (ATS) system have been previously enumerated in Section 3.3. The main functions, however, are repeated here for convenience and completeness:

1/ Data Acquisition

(operator records, vehicle occupancy, distance measurements, line and route information, etc.)

2/ Data Transmission

(cyclical interrogation of trains, information exchange, failure recognition)

3/ Data Evaluation

(reports, schedule comparison, travel reports, statistics, etc.)

4/ Controls

(Communications, passenger information, etc.)

Three system elements participate in the ATS functions:

1/ the train

2/ the wayside (position, station displays, etc.)

3/ the control (supervision) (Figure 28)

6.4.1 ATS Using Radio Communications

In the past, the most commonly used ATS systems were simple radio communication links between the drivers and the central communication exchange. This was done only in rare and abnormal situations, however, because of the insufficient number of frequencies which could be allocated to transit. Other operational information, such as deviations from schedule, number of passengers on board (loading conditions), vehicle status, etc., was usually not forwarded

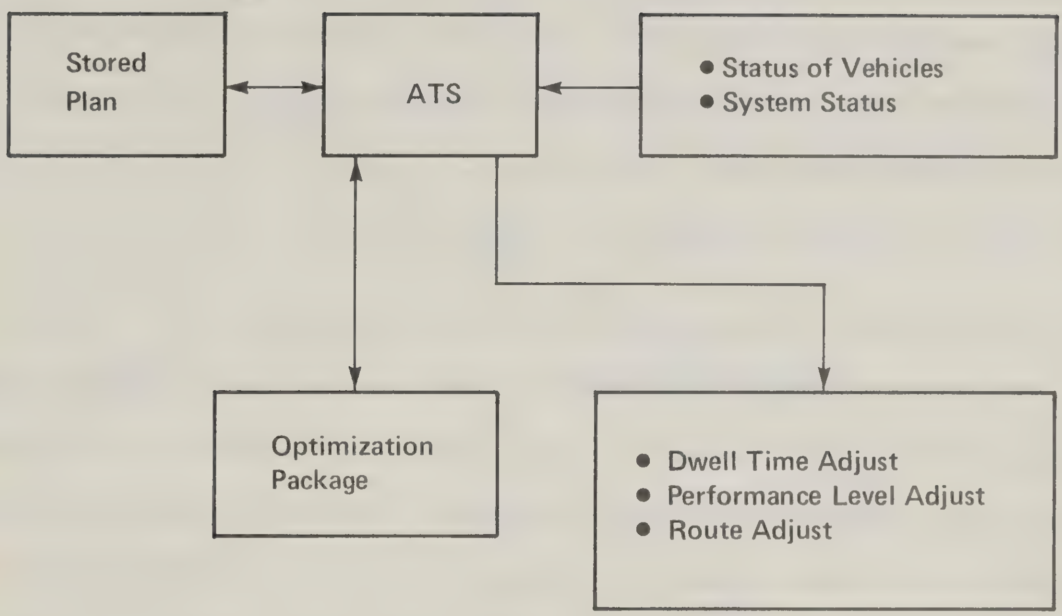


Figure 28; Control Supervision Under ATS

to the central for two reasons:

1/ The transmission of additional operational information would quickly exceed the available frequency channels.

2/ Voice communication is ill suited for processing and evaluating transit operations.

As Figure 29 shows, the simplest ATS is a voice communication link between the vehicles and the central where the information provided by the drivers is collected and evaluated and appropriate counter-measures are formulated. Due to the aforementioned system limitations, this type of ATS system cannot be used where very high operational efficiency is a goal.

6.4.2 Computer-Based ATS with Digital Communication Link

The need to improve operations in urban transit led to the development of computer-based, fully automatic train supervision systems. Full automation is restricted, however, to normal situations. The human operator is retained to cope with emergency situations, during which voice communication with drivers who need assistance or advice is particularly important. The fully-automated ATS system automatically keeps track of the position and status of each vehicle. System performance is compared to either a set of stored programs or optimized according to some optimization algorithms. In either case, a set of recommended actions are input to the ATS system. These "recommendations" are checked for safety and consistency before transmission to to the vehicles. Vehicle status conditions are also considered in formulating the "best" operating strategy.

In cases where the vehicles are operated by a human operator, the transmission of frequently recurring data is by push buttons; otherwise it is fully automatic. In either case, voice communication on request can be provided between the

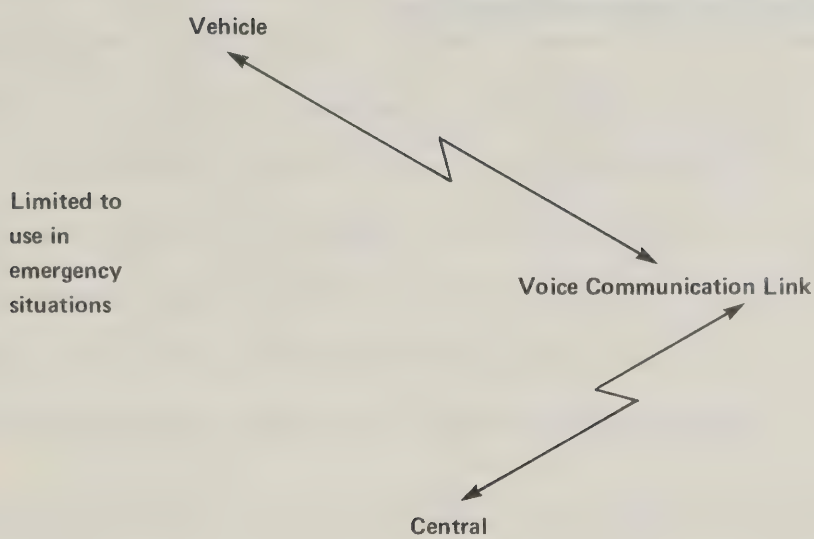


Figure 29; Simplest ATS — A Voice Communication Link Between Vehicles and Central

vehicles (passenger or driver) and central.

Between "fully" automatic ATS and voice communication dependent systems, there are various levels of train supervision. The "train describer", for example, is a device for transmitting a coded description of a train from location to location. It can be connected to the central control and used to provide limited information for the formulation of efficient operating strategies. ATS algorithms are not yet fully developed and operational experience to this time does not reflect the expected benefits.

7. COST/BENEFITS OF AUTOMATION

7.1 General

Previous sections of this report, have described various levels of automatic train control, their functional requirements, their structure, operational characteristics, limitations, and comparative advantages. At the train protection level, the need for automation is well recognized. The application of automation to train protection results in increased safety and also permits higher speeds of operation. The benefits of ATP, therefore, will be further examined. Automating the train operation and train supervision functions, however, has not yet been widely accepted, and careful analysis is required to determine the feasibility, the economic justification, and the overall appropriateness of the automation of these functions. The following section will deal with ATO benefits only since insufficient data exists to examine the benefits of various ATS algorithms.

Prior to undertaking such an analysis, a few ground rules must be established. The first, and most important assumption is that the overall transit system is designed (or redesigned, as the case may be) with the automation of the appropriate ATC functions in mind. Only if the cost-effectiveness of the whole system proves advantageous is automation introduced. To be specific, if the cost/benefit analysis shows that for ATO to be cost-effective, a certain minimum level of ATO equipment reliability has to be achieved, the technology and the cost of achieving the necessary reliability level must be included in the analysis. It does not appear reasonable to include ATO with a lower than adequate reliability figure only to conclude at a later stage that ATO does not prove to be cost beneficial. In the same vein, it appears unrealistic to specify ATO with the aim of reducing the train crew

only to find later on that union rules or other constraints do not permit the saving in cost to occur.

The second ground rule, or assumption which serves as a corollary to the first is that once a level of automation is specified, that level is also made use of. If the level of automation is specified as an option for later service, then the "value" of the option must be considered as a benefit.

In summary, the potential of each of the levels of automation must be established. The potential must be tempered by the other system constraints and the benefits of automation for each of the various systems established. This approach will suggest the conditions under which automation levels are feasible and justifiable and also the conditions under which the introduction of automation does not prove to be beneficial.

The following analysis will distinguish between three cases:

1/ The introduction of ATO does not reduce the train crew.

2/ The introduction of ATO reduces the train crew by 1 without eliminating the last crew member.

3/ The introduction of ATO results in unmanned train operations.

All of the three cases are related back to manual operations. Since Automatic Train Supervision algorithms are not sufficiently developed, an analysis of ATS benefits will be omitted; and only the benefits of ATO will be analysed.

7.2 Increased Capacity

Automatic Train Operation (ATO) has the potential to yield higher system capacity. The potential for higher capacity is in itself a benefit but,

more realistically, this capability should only be considered a true benefit if higher capacities than achievable by manually operated systems are required and if these higher capacities are actually achievable in the system, i.e., if they are not limited by physical constraints. These constraints take on different forms depending on whether the trains are manned by one or more operators, or whether the trains operate unmanned. The operating and safety policies and the physical layout of the system also exert an influence. At first, the potential for capacity gain through the use of ATO will be approximated and then the constraining factors are introduced to indicate when, and under what conditions the capacity gain can be achieved.

For a fixed-block system, it can be shown that ATO can reduce line headways by approximately 4 s to 6 s relative to manual systems. This headway gain arises from the improved speed and acceleration/deceleration regulation that can be achieved by the ATO system.

It is common practice to assume a ±10% regulation tolerance band on the speed and acceleration/deceleration of manually controlled trains. These assumptions lead to a system k-factor of 1.35 (k = worst case stopping distance/nominal stop distance). Automatic Train Operation would reduce the speed regulation accuracy to 3% and the acceleration/deceleration band to approximately ±7%. For these figures, the k-factor with ATO becomes 1.14. This reduction in k-factor can result in a roughly 4 s headway improvement.

If a moving-block type system is used for headway regulation, it can be shown (Appendix A) that a potential station through headway improvement of 10 s to 15 s with respect to manual fixed-block operation can be obtained.

This brief analysis yields the approximate limits of the potential capacity increase with ATO. Before accepting these limits, some extraneous conditions must be considered to decide whether ATO can, or cannot realize its potential for increased capacity.

7.2.1 Turn-Around Time

In many rapid transit systems with on-line stations the turn-around time at the end terminals becomes the limiting factor on the achievable headway. The turn-around time, however, is limited by physical factors (e.g., switch area geometry) and safety policy as well as the vehicle characteristics. ATO would not significantly improve the system headways which are limited by the turn-around time. If a 2-man operation with ATO is considered, and the guard is permitted to walk through the train before the end terminal has been reached to take his place as the driver in the reverse operating direction, ATO could lead to a 2 to 3 s headway reduction.

This improvement would be largely a result of shorter reaction times with ATO than possible in manually operated systems. (This also assumes the start signal is not device activated). If ATO resulted in a 1-man train operation, the turn-around time and, consequently, the system capacity could be improved or worsened depending on the manner in which the turn-around of the trains is effected. With a crossover turn-around the driver would have to walk from one end of the train to the other causing significant delays A "roving" driver at the end terminal would eliminate this delay and the gain in headway would be limited to 2 to 3 s <u>if ATO was responsible for starting the vehicles</u>. Unmanned train operation would likely result in a 2 to 3 s headway gain at the turn-around.

7.2.2 Through Station Capacity

In many urban transit systems, the headways required at the end terminal turn-around are substantially greater than those through the busy sections of the downtown area. To achieve shorter headways in the busy central sections, vehicles are either short-turned (in which case the limitations discussed in Section 7.2.1 still apply) or funneled from many directions (branches) into a central route section (Trans-bay tube in San Francisco, downtown area Munich, etc.). If no short-turning of the vehicles is required in the busy central section of the route, ATO potential for increased capacity can be realized and a 10 to 15 s headway gain with respect to manually operated vehicles over the same section can be achieved. This capacity increase can be achieved independently of any crew considerations. The headway through the central section is then limited only by the station through headway achieved in that section (or outside of it) and not by the usually more constraining turn-around.

7.2.3 Vehicle Door Control Procedures

Vehicle door control procedure also has a bearing on whether ATO can realize some gain in system capacity. In 2-man operation, the guard usually controls the closing of the vehicle doors. In 1-man operation, the driver is often required to step out of his cab and supervise the safe loading of the passengers. In some systems, and certainly in unmanned train operation, the door opening and closing must be fully automatic. It is unlikely that automatic door closures can be as efficient in terms of station dwell as systems employing 2 men. Door cycle time can only be set to some mean value and in many cases will either be greater than required or less than required. The station dwell times under automatic door control will be

longer than under manual door control and this will lead to longer headways and lower system capacities.

7.2.4 Summary

It was shown that the introduction of ATO to rapid transit system application results in a potential capacity increase. Whether this potential is realized or not depends on a large number of system specific factors which must be considered very early in design of the total transit system. In some applications, the potential gain in capacity is unimportant and no direct cost benefits can be assigned to this particular ATO capability. For systems which operate close to capacity and which are not limited by other system factors, such as turn-around time, door operations, etc., a reduction of up to 15 s in the system headway can be very significant.

7.3 Smaller Stations (Shorter Train Consists) WITH Automation

Proponents of automation have repeatedly stated that automation leads to shorter trains and consequently shorter stations. The reason quoted for the reduction is the shorter headway which automated trains can achieve. The following analysis will show that automation does not always lead to shorter trains and shorter stations. The potential saving in station length is readily determined by first examining the number of vehicles in a train N_v required to carry an hourly passenger demand C at a headway of H seconds and with a vehicle capacity f_t passengers per vehicle:

$$N_v = \frac{C \times H}{f_t \times 3600} \text{ vehicles/train}$$

If H_m denotes the headway achievable under normal manual control, H_a the headway achievable with automation, N_{vm} the number of vehicles per train in manual operations, and N_{va} the number of vehicles per train for automatic

operation, the following simple expression can be written for constant demand and vehicle size:

$$\frac{N_{va}}{N_{vm}} = \frac{H_a}{H_m}$$

or $$N_{va} = N_{vm} \times \frac{H_a}{H_m}$$

The above expression indicates the potentially achievable reduction in the number of vehicles per train if a system is automated. Since the station platform length is roughly equal to the maximum train size, the corresponding relation for the potential station reduction with automation is:

$$l_a = l_m \times \frac{H_a}{H_m}$$

Simply stated, the percentage reduction in train or station length due to automation is equal to the percentage reduction in the headway.

The above expression assumes implicitly that train sizes can be continuously adjusted. Trains, however, are made up of vehicle units (single vehicles or married pairs) of finite length. Fractional vehicle units do not exist so train length must be varied by adding or removing vehicle units. Also, continual adjustment of train length leads to significant operational difficulties.

In Section 7.2.4, it was concluded that an approximate 15 s reduction in headway is possible by using automatic train operation. Table 14 assumes that a 10 s reduction (typical value) in headway has been achieved by automatic train operation.

Interpreting Table 14, the following general conclusions can be reached about

Table 14; Train Consist vs. Headway
(Under Manual and Automatic Control)

DEMAND (pphpd)	MODE	HEADWAY ACHIEVABLE (s)	No. of Vehicles Req'd to Carry Stated Demand		
			VEHICLE CAPACITY 50 psg/veh	VEHICLE CAPACITY 100 psg/veh	VEHICLE CAPACITY 200 psg/veh
20 000	Manual	60	6	4	2
	Automatic	50	5	3	2
	Manual	120	14	7	4
	Automatic	110	13	7	4
40 000	Manual	60	14	7	4
	Automatic	50	12	6	3
	Manual	120	27	14	7
	Automatic	110	25	13	7

the practicability of train and station size reduction by automation:

1/ Systems using smaller vehicles are more likely to make use of the 10 s headway reduction through automation resulting in shorter trains than would be possible with manual control. However, in the more unlikely case that a headway reduction eliminates the need for a larger vehicle unit, the unit gain is higher.

2/ The higher the peak demand, the more likely that automated trains will run with shorter consists than their manually-operated counterparts.

3/ The shorter the peak-hour headway, the more likely automation will result in shorter train consists than manual train operation.

4/ Only in new systems is a reduction in station size possible. Existing systems operate with station sizes which have been designed with manual operation in mind. As illustrated in Table 14, even in new systems, it is not always possible to translate the headway reduction into a station size reduction.

5/ The headway reduction due to automation is subject to the constraints outlined in Section 7.2

7.3.1 Service Frequency

From a passenger's point-of-view, a 15 s reduction in headway which may potentially be achieved with ATO (see Section 7.2) is not perceived as being very significant. If, however, analysis could show that systems with ATO are cheaper to operate than manually-operated systems, the difference in operating costs could be used to increase the service frequency in off-peak hours. The following general statements may be made at this point:

1/ If the introduction of ATO does not result in the removal of at least one crew member from the train, then no increase in off-peak service

frequency can be offered to the transit riders (over and above that offered by manually-operated systems) without a corresponding increase in operational cost.

2/ Unmanned operation of trains allows high off-peak service frequencies without an increase in operating cost over and above the base cost.

ATO offers the potential for higher service frequency in off-peak hours but this potential exists only if the cost of a system run under ATO control is lower than the cost of a manually-operated system. (Later sections discussing ATO costs will focus on this question.) The service frequency offered is also a function of the system reliability. The ATO controlled systems must be at least as reliable as manually-operated systems in order to offer equal or higher service frequencies.

In summary, ATO introduction which results in a train crew reduction of at least one man, but ATO operation with unmanned trains offers the greatest potential of high frequency service to the transit patrons.

7.3.2 Punctuality of Service

"Punctuality of service", or "schedule adherence" is a complex function which depends on many terms and which may be defined in many different ways. A recent report by the U.S. Office of Technology Assessment* quotes an overall index of schedule adherence for the PATCO (Philadelphia Port Authority Transit Corporation) system as:

$$\text{Schedule Adherence} = 100 \frac{(T_s - T_a - T_l - 0.1 \times S_b)}{T_s}$$

---

*Automatic Train Control in Rapid Transit Systems, OTA, May 1976.

where T_s = trips scheduled

T_a = trips annulled

T_l = trips late

S_b = stations bypassed

The schedule adherence in PATCO for 1974 is quoted in the same report as:

Performance, 1974

(Scheduled Trips)	
Percent on time	98.36
Percent late	1.16
Percent annulled	0.48
Stops bypassed in %	0.18
Calculated schedule adherence in %	98.34

For the Chicago Transit Authority (CTA) the same report undertook a comparative analysis with the following performance parameters:

Performance, 1974	
Percent on time	96.26
Percent late	3.65
Percent annulled	0.09
Stops bypassed in %	0.34
Calculated schedule adherence in %	95.92

The report went on to conclude that "it would appear that a manual system with ATP (CTA) and an automated system with ATP and ATO (PATCO) can achieve equal levels of schedule adherence". Since advocates of ATO often quote better schedule adherence as one of the ATO advantages, these findings

deserve a close evaluation. Firstly, if one calculated the schedule adherence according to the index provided for both systems, PATCO would show a schedule adherence of 98.34% and CTA 95.92%. This in itself would indicate a 2.4% better schedule adherence in the PATCO system. Examination of the elements which made up the schedule adherence index, however, proves to be very revealing.

Only 1.16% of the trips scheduled were late (defined as a late arrival at the terminal of 5 min or more) in the case of the PATCO system as compared to 3.65% for CTA. The percent of annulled trips was 0.48% for PATCO and 0.09% for CTA. Bypassing of scheduled stops was 0.18% for PATCO and 0.34% for CTA. Since a very large number of factors contribute to delays in a system, it is difficult to draw firm conclusions, nevertheless it appears that the system with ATO (PATCO) was roughly 3 times more effective in keeping late arrivals below 5 min than the only ATP-equipped CTA system. Presumably, due to the lower reliability of the PATCO system, the percent annulled trips was higher than for CTA. The differences may be small but they seem to support the claim that systems with ATO are capable of better schedule keeping than systems under manual control. At this time, however, it appears that the full capability of such systems is overshadowed by the reliability problems which they experience.

The experiences on the Victoria Line in London seem also to support the better schedule keeping ability of ATO. Tests conducted on this line showed an average running time variation between stations of ±3.1 s. The variations on the manually-run Piccadilly Line were ±4.6 s. The improved schedule adherence can be directly translated into capacity gain and/or fleet size reduction and, based on the results from PATCO and the Victoria Line, this

better schedule adherence is assumed to translate into a capacity gain of 1.5%.

In summary, it appears that systems with ATO do show better schedule adherence than manually-operated systems. The differences are small but significant. Based on the operating experience of two systems, it is estimated that the better schedule adherence of ATO systems results in an approximately 1.5% capacity gain over manually-controlled systems. This capacity gain arises from the more even distribution of vehicles along the route and the corresponding improvement in vehicle loading.

7.3.3 Travel Speed

The increase in travel speed does not appear to be very significant and will not be further considered.

7.3.4 Transit Personnel

The hiring and training of good transit operators is a difficult task. Automatic Train Operation reduces the problem of hiring, training, and retaining the appropriate personnel and also minimizes the risk of inadequate performance. If unmanned operation of trains is permitted,the driver problem is totally eliminated.

It is claimed that ATO eliminates the driver problem but substitutes for it the need for skilled technical people. There are two sides to this statement: the present problem of the transit authorities who lack the skilled technical people on their staff to operate and maintain the more complex ATO equipment, and the future problem of adequate supply of people skilled enough to work and repair the ATO equipment. The first appears to be a serious but transient problem, the second is unlikely to become a real problem since the supply of highly-trained technical people appears to be assured.

The question of relative costs, i.e., the cost of the operators versus the cost of the technologists is, however, important and will be dealt with in Section 7.3.6.

7.3.5 Energy Consumption

One advantage claimed for ATO (with ATS) is the ability to keep schedule with minimum energy use, i.e., traction cost. The interstation running time is largely governed by the cutout speed, i.e., the maximum speed that is reached during the acceleration phase. The most energy efficient run is one in which trains only accelerate to that speed which is required to enable the train to arrive on schedule at the next station. Such preprogrammed running can aid in good schedule adherence with minimum traction cost. Manually-operated systems can also be programmed but the operating control and accurate decision-making is best left to automated systems.

Since automated systems continuously collect all headway, and other operational information, the acquired information can be used to optimize the energy consumption of the trains. ATO affects the energy savings in 3 ways:

1/ ATO under ATS control follows a "schedule keeping with minimum energy consumption" strategy.

2/ More accurate regulation of cutout and brake points is possible.

3/ ATO achieves more uniform headways resulting in more uniform braking and acceleration. The more uniform spacing of the vehicle also has an equalizing effect on the passenger load of trains resulting in more energy savings.

Appendix B, Table B1, shows the theoretical energy savings with ATO. These

savings, however, are unlikely to materialize in full. If the transit authority operated under the policy of "maximum performance under peak conditions", there would be no way of achieving energy savings by trying to control the cutout point since the cutout point under these circumstances corresponds to the maximum speed. If unmanned operations were in force, the door cycle time would have to be set to some average value. This procedure would effectively eliminate most energy savings which would be possible if the trains could depart before the average door cycle time had expired.

In summary, some energy savings are possible under ATO-controlled operations. The magnitude of these savings depends to a large degree on system specific factors such as door operating procedures, manned or unmanned train operations, the existing operating policies, and vehicle characteristics. Also, the savings represent only fractions of a percent of the total operating cost.

7.3.6 Operating Cost Reduction

The reduction in operating costs projected for ATO originates from:

1/ Reduction of the crew

2/ Reduced vehicle wear and tear

3/ Energy cost reduction

These cost reductions due to ATO must be appropriately weighed against the corresponding cost increases. ATO may allow a single person to operate vehicles and result in a lower driver payroll. Against this gain, one must weigh the increased maintenance cost and the other associated expenditures. The energy cost savings due to ATO have been discussed in Section 7.3.5 and were shown to be very system specific. Reduced vehicle wear and tear is

unlikely to become a very important factor and too many factors contribute to wear and tear to isolate the importance of ATO. The largest single reason for ATO cost reduction is the possibility of train crew reduction from 2 or more to 1 man only and then to unmanned train operations. It must be emphasized that automation is only one of many factors contributing to the size of a train crew. Safety, security, union regulations, work rules and other system specific factors contribute to the train crew size definition.

As in all cases discussed in this chapter, it is most important to determine the cost/benefit potential of automation first, then superimpose constraints and limitations which tend to reduce the benefits. In the following, a very simplified cost analysis for ATO is presented. To the extent possible, the analysis uses representative values and assumptions and the intent is to show trends rather than absolute magnitudes.

7.3.6.1 Cost Analysis for ATO - Personnel Costs

The following assumptions are used in the cost analysis:

1/	Driver cost	= $10/h
2/	Technical maintenance personnel cost	= $12/h
3/	ATO on-board equipment mean time between failures (MTBF)	= 1000 h
4/	ATO wayside equipment MTBF	= 500/h/km (single track)*
5/	ATO on-board mean time to repair (MTTR)	= 3 h
6/	ATO wayside MTTR	= 3 h
7/	Mean interruption time due to ATO failure (without operator on-board)	= 5 min

*Derived from BART and PATCO cost figures

8/ Mean interruption time due to ATO failure (with driver on-board) = 0 min

9/ Cost of on-board ATO equipment (incremental cost*) = \$6000/vehicle**

10/ Cost of ATO test equipment = \$1.0/vehicle

11/ Cost of ATO wayside equipment (incremental cost*) = \$40,000/km (single track)**

12/ Peak hours/day = 4 h

13/ Total operating hours/day = 20 h

14/ Minimum scheduled headway = 60 s

15/ Operating speed = 20 m/s

16/ Peak demand = 40 000

17/ Off-peak demand = 4 000

7.3.6.1.1 Approximate Cost of Manual Train Operation

Taking "n" to be the number of peak-hour trains, "X" the ratio of peak-hour to be off-peak hour trains, "G" the hourly cost of drivers and "p" the number of operators per train, the cost of manually driving the trains can be written as:

$$C_p = \frac{1}{n}\left[4\left(\frac{\text{peak-hours}}{\text{day}}\right) \times p\left(\frac{\text{driver}}{\text{train}}\right) \times n \text{ (number of peak-hour trains)} \times G\left(\frac{\$}{\text{driver-hour}}\right) + 16\left(\frac{\text{off-peak hours}}{\text{day}}\right) \times p\left(\frac{\text{drivers}}{\text{train}}\right) \times n \text{ (number of peak-hour trains)} \times \frac{1}{X}\left(\frac{\text{off-peak trains}}{\text{peak trains}}\right) \times G\left(\frac{\$}{\text{driver-hour}}\right)\right]$$

$$= G \times p \times 4\left[1 + \frac{4}{X}\right] \text{\$/peak-hour train day.}$$

---

*Incremental costs are with respect to a baseline ATC system consisting of Cab Signalling, Overspeed Protection, and Route Interlocking.

**Derived from BART and PATCO cost figures

Substituting \$10/h for G, the cost of driving the trains is shown in Figure 30 for one and two operators per train (p = 1 and 2). The daily driving costs per peak-hour train are drawn as a function of X, the ratio of peak to off-peak hour trains. It can be seen that the cost of driving is highest if equal service is maintained in off-peak hours (X = 1) and decreases with decreasing service level (X).

7.3.6.1.2 ATO-Related Costs

7.3.6.1.2.1 ATO-Equipment Cost

The ATO equipment cost consists of the vehicle-based equipment cost plus the cost of the wayside equipment and the central computer. It is assumed that automatic train protection with cab signalling is already part of the system and that only the incremental cost of upgrading the system to ATO must be considered.

The wayside equipment cost can be translated into vehicle-based costs in the following manner:

$$\text{ATO wayside cost } [\$/\text{vehicle}] = \frac{H \cdot v \cdot C_w}{N}$$

where H = minimum peak-hour scheduled headway (s/train)

v = schedule velocity (m/s)

C_w = ATO wayside cost (\$/ m (single track)

N = number of peak-hour vehicles/train

The total incremental ATO equipment cost is calculated by adding the on-board ATO equipment cost to the wayside ATO equipment cost:

$$C_A = 6000 + \frac{H \cdot v \cdot C_w}{N} = 6000 + \frac{H \cdot 745 \cdot 8}{N} \quad (\$/\text{vehicle})$$

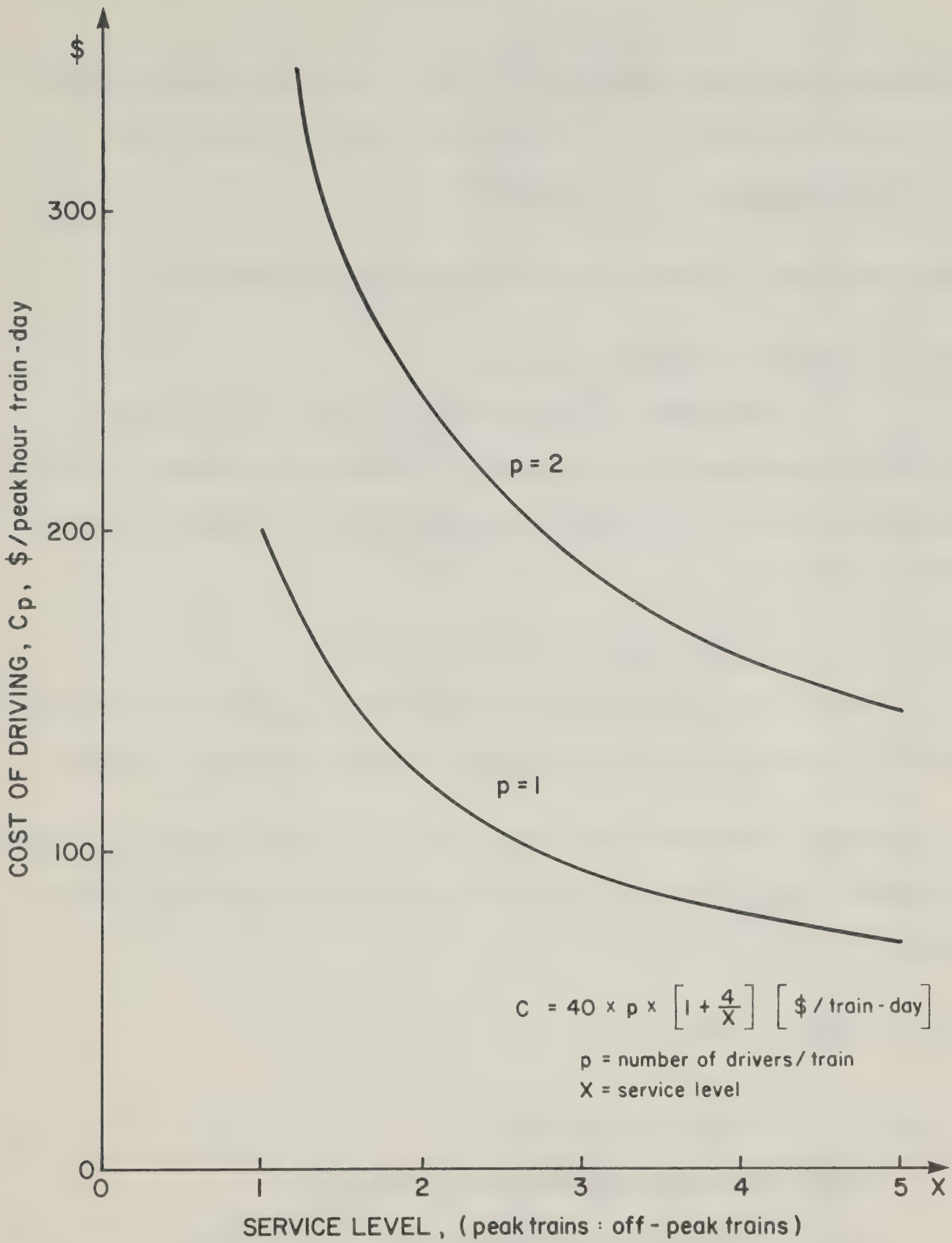


Figure 30; Cost of Driving a Train vs. Service Level

Considering an amortization period of 20 years and an amortization rate of 12%, the daily incremental ATO equipment costs can be calculated as:

$$C_D \text{ (\$/vehicle-day)} = 1.9 + \frac{H \times 0.245}{N}$$

Figure 31 depicts C_D as function of minimum scheduled headway, H.

7.3.6.1.2.2 ATO Test Equipment Cost (C_E)

It is reasonable to assume that additional test and repair facilities and equipment will be required to support the automatic train operation (ATO) system. An amortized figure of $1/vehicle-day is assumed to reflect this cost.

7.3.6.1.2.3 Cost of ATO Repair and Maintenance (C_M)

The cost of ATO repair and maintenance consists of two elements: the cost of on-board failures and the cost of wayside equipment failures.

As a simplifying assumption, the failure rate of the ATO equipment is assumed to be constant over 24 h. The cost due to on-board failures can then be calculated:

$$C_{BA} = \frac{t_d}{MTBF_o} \cdot MTTR_o \cdot C_o$$

where t_d = hours per day (h/d)

$MTBF_o$ = on-board ATO equipment MTBF/vehicle

$MTTR_o$ = mean time to repair, on-board ATO equipment

C_o = Cost of repair ($/h)

Substituting for the parameters, using C_o = $12/h, $MTTR_o$ = 3 h, t_d = 24 h, $MTBF_o$ = 1000 h/vehicle, the cost of on-board ATO failures becomes:

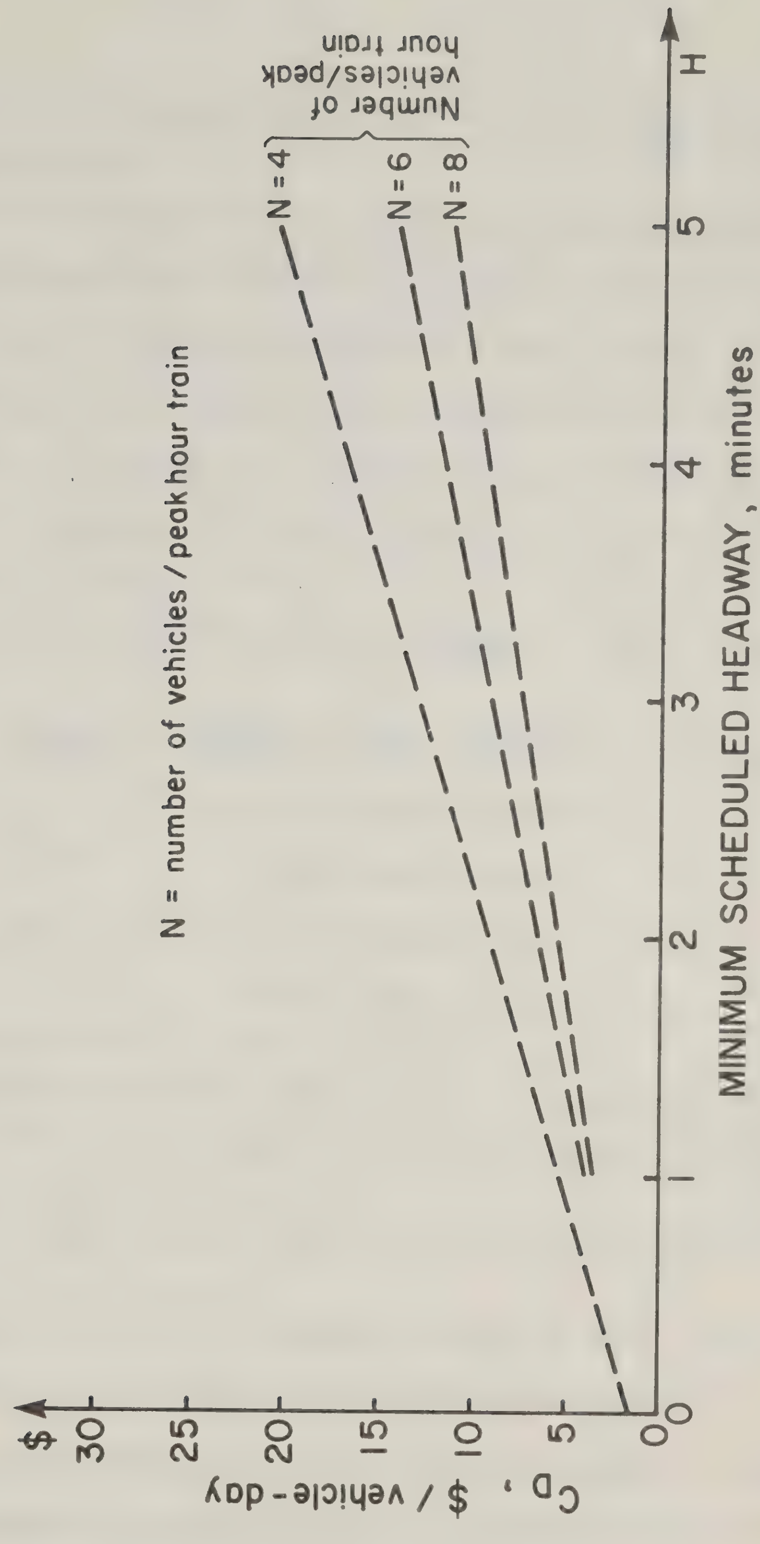


Figure 31; Incremental ATO Equipment Cost (C_D) as a Function of Minimum Scheduled Headway (H)

$$C_{BA} = \frac{24}{1000} \times 3 \times 12 = \$0.86/\text{vehicle-day}$$

The wayside ATO equipment failure may be transformed into a vehicle-based failure to facilitate comparison. Facility costs, spare part costs are not separately shown. Since the schedule velocity of 20 m/s is considerably higher than normally achieved, all calculated costs are higher than expected and are therefore assumed to contain the additional costs of overhead, spare parts, etc. The transformation, then, is as follows:

$$C_{WA} = \frac{H \cdot v}{N} \cdot \frac{1}{MTBF_w} \cdot t_d \cdot MTTR \cdot C_x$$

where H = minimum scheduled headway (s/train)

v = schedule velocity (m/s)

N = vehicles/peak-train

$MTBF_w$ = wayside MTBF in h/m (track)/failure

C_x = cost to repair ($/h)

Using the figures of v = 20 m/s, $MTBF_w$ = 500 000 h/failure/m, t_d = 24 h, the total cost of ATO failures can be written as:

$$C_M = C_{BA} + C_{WA} = .86 + 0.034 \frac{H}{N} \text{ (\$/vehicle-day)}$$

7.3.6.1.2.4 Cost of Service Interruptions due to ATO Failures (C_s)

Two cases must be distinguished. In the first case, ATO operation is supplemented by at least one operator who, if necessary, can drive the vehicle manually. For this case, it is assumed that an ATO failure does not result in a service interruption. The second case, that of unmanned operation, results in a delay for every ATO failure. To keep this delay to a mean of 5 min or less, a "roving maintenance crew" is required

consisting of 1 man for every 20 vehicles*. Also, additional siding installations are postulated.

The cost of a siding will, of course, depend on the system (whether, for instance, it is at-grade, elevated, or underground). It is also very difficult to assess, without a detailed analysis, how many additional sidings are required, however, as a first approximation a cost of $50,000 per additional siding will be used. The placement of these sidings (additional to the number required in manual operations) will be every 10 000 m. The cost per vehicle of the additional sidings is, therefore:

$$50\ 000 \frac{\$}{\text{siding}} \cdot \frac{1}{10\ 000} \cdot \frac{\text{siding}}{\text{meter}} \cdot \frac{H \cdot v}{N} \quad \frac{\text{meter}}{\text{vehicle}}$$

$$= 5 \ \frac{H \cdot v}{N} \quad \frac{\$}{\text{vehicle}}$$

$$= 0.033 \frac{H}{N} \ \$/\text{vehicle-day}$$

This is the amortized cost over 20 years at 12%/a.

Neglecting any time losses to passengers, the total service interruption cost reduced to a per vehicle-day basis is:

$$C_s = 1/20 \ t_d \cdot C + 0.033 \frac{H}{N}$$

$$C_s = 1/20 \cdot 20 \cdot 12 + 0.033 \frac{H}{N} = 12 + 0.033 \frac{H}{N} \ (\$/\text{vehicle-day})$$

where 1/20 = number of maintenance men/vehicle

C = cost of maintenance man ($/h)

---

*This ratio is based on PATCO and BART experience

The above expression for C_s, the cost of service interruption, has been derived on the basis of equal service frequency for peak hours and off-peak hours during the day. To be more accurate, the roving crew requirement must be adjusted according to the time of the day, since it is expected that the off-peak demand will be lower than the peak demand and consequently fewer vehicles will be required to meet the off-peak demand. Since the ratio of off-peak, to peak demand was assumed to be 4000 : 40 000, (i.e., in a ratio of 1 :10), a corresponding reduction in the number of off-peak vehicles appears justified. To assure seats for all off-peak riders, a 1 : 4 instead of a 1 : 10 vehicle ratio is assumed.

The adjusted service interruption cost C'_s then becomes:

$$C'_s = (1/20 \cdot N \cdot 4 \cdot 12 \quad + 1/20 \cdot 16 \cdot 12) \cdot \frac{1}{N} + 0.033 \frac{H}{N}$$

$$= 4.8 + 0.033 \frac{H}{N} \quad (\$/\text{vehicle-day})$$

Lost passenger time cost due to interruptions has not been included in these calculations because its inclusion in the analysis would probably overshadow all other trade-offs. If required, the methodology can readily be extended to include the cost of passenger time lost due to service interruptions.

7.3.6.1.2.5 Total ATO-Related Cost

The total ATO related costs may be summed up as:

C_T = ATO Equipment Cost (C_D) + ATO test equipment cost (C_E) + ATO repair and maintenance cost (C_M) + ATO service interruption cost (C_s*)

---

*The term C_s is used when the service frequency in peak and off-peak hours is assumed to be equal. Otherwise, the term C'_s, may be used. It represents adjusted service interruption cost in which service frequency is adjusted to demand by a change in the number of vehicles.

$$C_t = 1.97 + \frac{0.245H}{N} + 1.0 + .86 + 0.034\,\frac{H}{N} + (12 + 0.033\,\frac{H}{N})$$

$$= 15.83 + \frac{0.312H}{N} \quad (\$/\text{vehicle-day})$$

If the number of off-peak vehicles is adjusted to the demand (1/4 of the number of peak-hour vehicles) then:

$$C_T = 8.63 + \frac{.312\ H}{N} \quad (\$/\text{vehicle-day})$$

The above dollar values have been derived under the assumption of unmanned operation. Service interruption costs are assumed to be zero if a back-up driver is retained on-board the train to manually drive the vehicles in case of ATO failures. The total ATO cost for manned operations is:

$$C_{TM} = 3.83 + \frac{.275\ H}{N} \quad (\$/\text{vehicle-day})$$

The above equation is plotted in Figure 32.

There is also a small cost difference between ATO equipment used with human back-up and ATO equipment with no back-up. The cost differential is mainly due to the lower MTBF required of ATO equipment used with human back-up since all ATO failures can be compensated for by the on-board driver. It is unlikely, however, that the cost differences are significantly great to alter the cost comparison. Consequently, the analysis does not account for these differences.

7.4 Discussion of the Analysis and the Results

While an attempt has been made to reflect accurately the various elements which make up the ATO-related costs, the complexity of the subject precludes the achievement of absolute results. The analysis is intended to show a method and to arrive at "ball-park" figures which are meant to

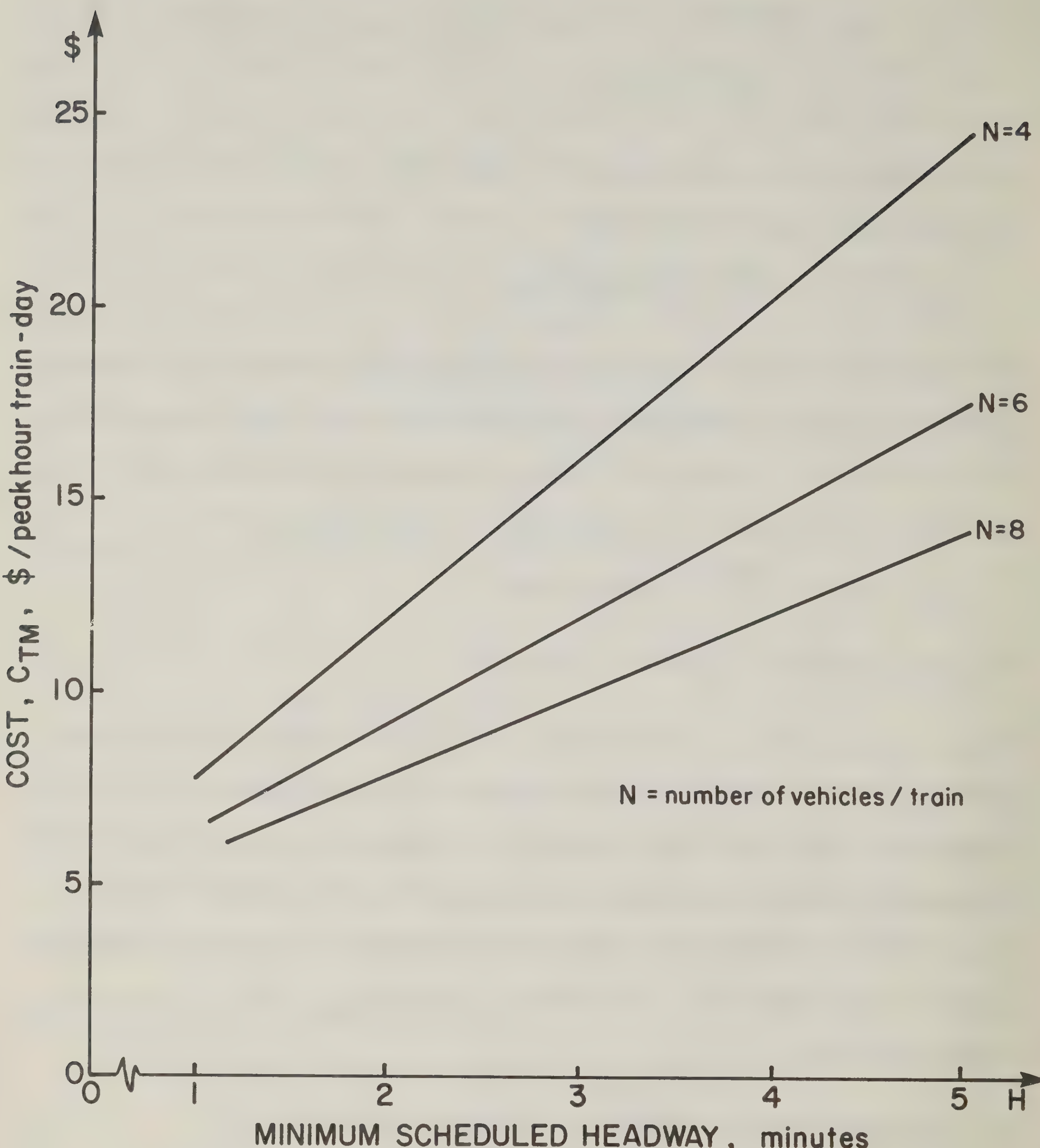


Figure 32; Cost of ATO with Manual Back-Up (C_{TM}) vs. Minimum Scheduled Headway (H)

show general trends rather than absolute values. The incremental cost required to upgrade a baseline system consisting of cab signalling and overspeed protection to automatic train operation is shown in Figures 31, 33 and 34 drawn as a function of headway. The baseline system of cab signalling with overspeed protection was chosen since it is considered to be the minimum ATC system for newly-implemented rapid transit systems (e.g. WMATA, MARTA)

As the above figures show, all ATO incremental costs, except for the test equipment cost C_E are functions of the minimum scheduled headway. The reason for this dependence is that wayside equipment costs were translated to equivalent vehicle costs thereby incurring a headway dependence. The translated costs per vehicle decrease with higher service frequencies since the fixed wayside costs can be apportioned to a larger number of vehicles (assuming a fixed capacity and speed). A substantial reduction in the costs of ATO-operated systems is achieved by reducing the off-peak train size to meet the required off-peak demand. In manually-operated systems, this is rarely done since driver costs are unaffected and also because coupling/ decoupling costs can become excessive. (In case of automatically operated systems, the actual cost of vehicle coupling and decoupling is very difficult to define. If the train size can be varied automatically and without incurring high operating costs, then the overall operating cost may be reduced by reducing the off-peak train size.)

Automatic systems which allow quick train formation gain in two related ways by reducing the off-peak train size. Firstly, there is a reliability gain for the non-automatic subsystems of the unused fleet and, secondly, there is a smaller likelihood of service interruptions due to failure of the

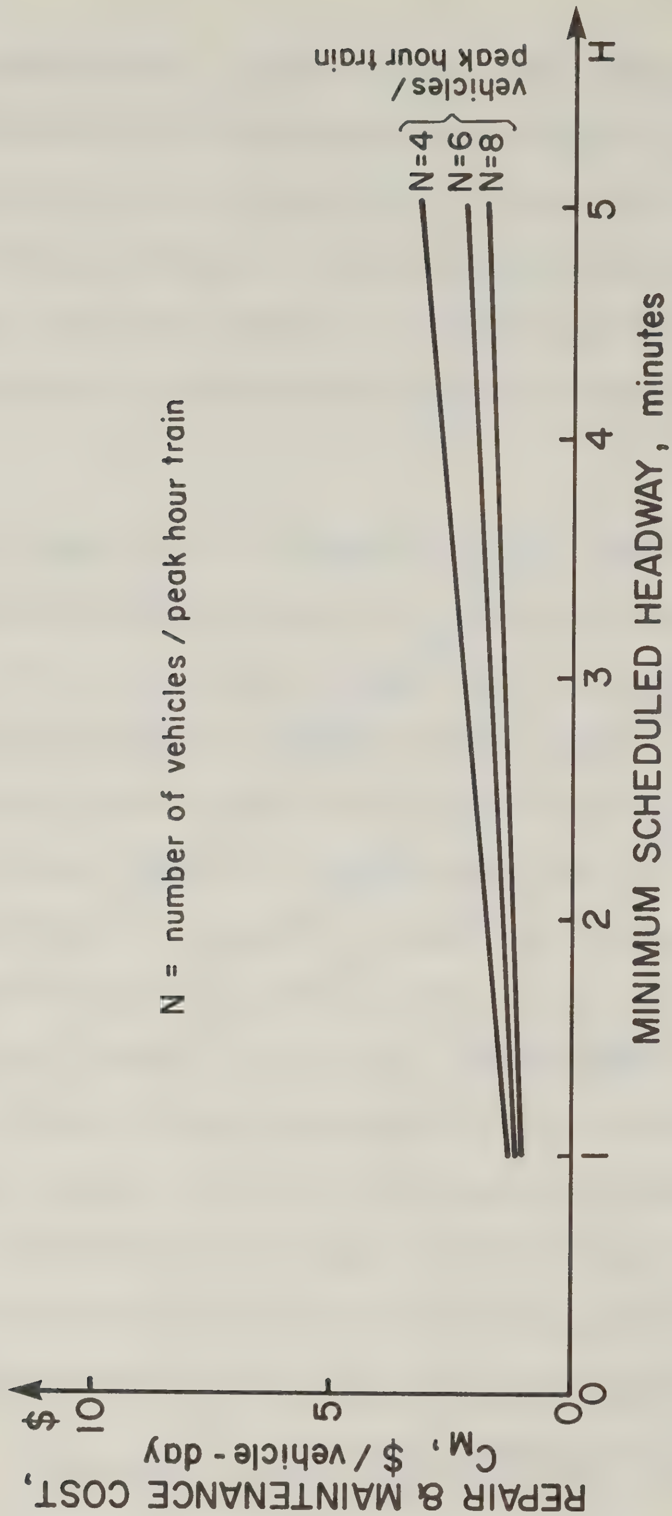


Figure 33; Incremental Cost of Repair and Maintenance (C_H) vs. Minimum Scheduled Headway (H)

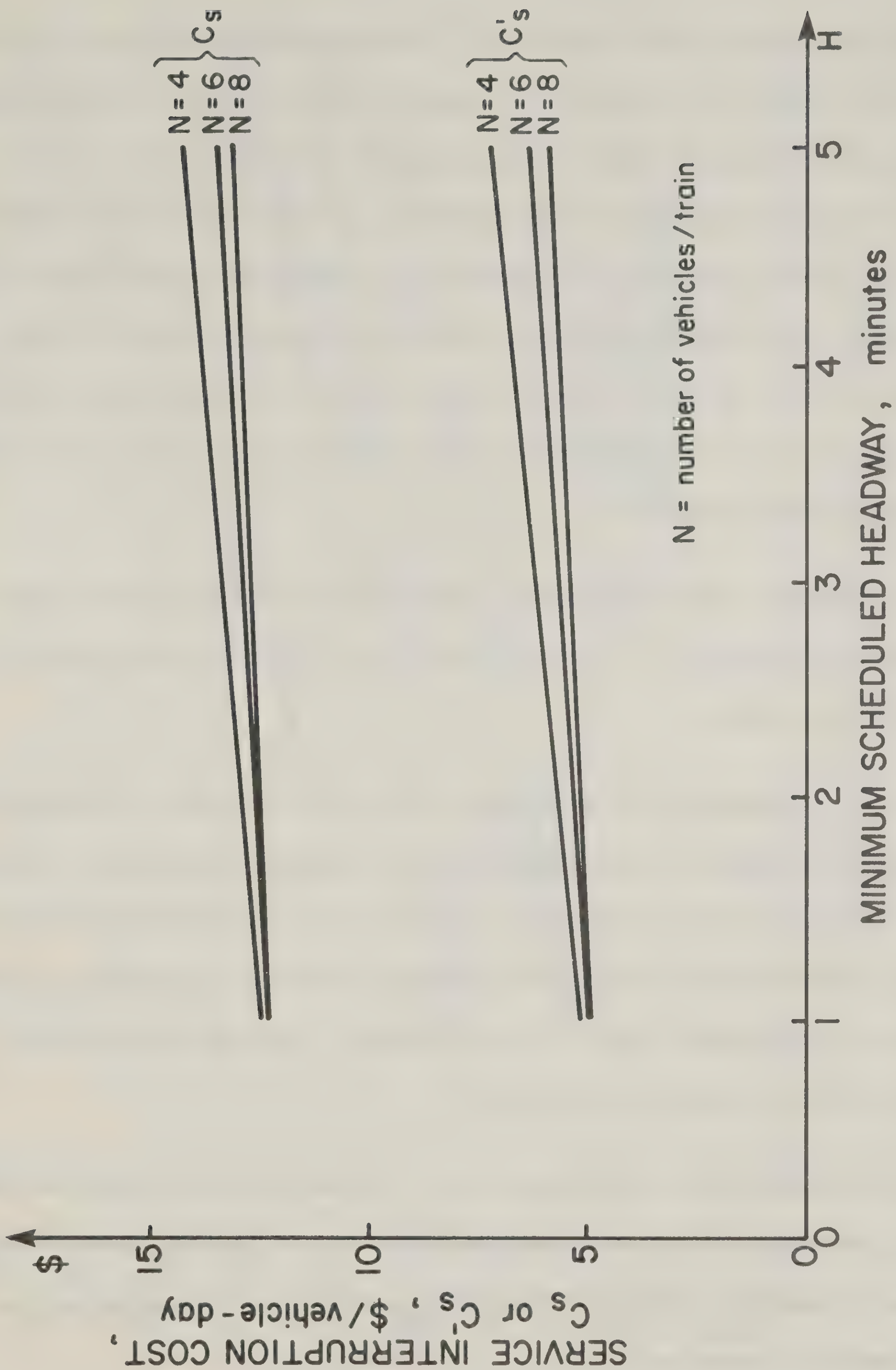


Figure 34; Unadjusted and Adjusted Service Interruption Costs (C_s & C'_s) vs. Minimum Scheduled Headway (H)

operational automatic (ATO) equipment. This reliability gain may not be fully or even partially realized if single cars are to be run at off-peak hours. Multiple-vehicle trains offer a higher reliability than single vehicles due to the redundancy they provide. If, however, married pairs are used and trains can only be augmented or reduced by complete married pairs, then reliability gains can be achieved by reducing the off-peak train size. Figure 34 gives a rough approximation of service interruption cost reductions which are possible by adjusting the off-peak train size to the reduced off-peak demand.

Figure 35 shows the total incremental ATO costs for both the case where off-peak trains are reduced, and for the case where they are kept intact for the whole operational day.

In Figure 36, a comparison is made between the cost of manual train operation and the cost of unmanned train operation, for various train sizes. Whereas the cost of manual train operations increases with increasing service level*, the cost of unmanned train operation remains fixed for all service levels. It is not very much more costly to run unmanned trains at peak service frequencies than at reduced frequencies.

As the driver productivity increases, i.e., as the number of vehicles per train, N, increases without a corresponding increase in the number of drivers, P, the cost comparison between manual and unmanned operations swings in favour of manual operation.

For rapid transit systems which normally operate at approximately 2 min peak

*Service level is defined as the reciprocal of the ratio of peak hour to off-peak hour trains (X).

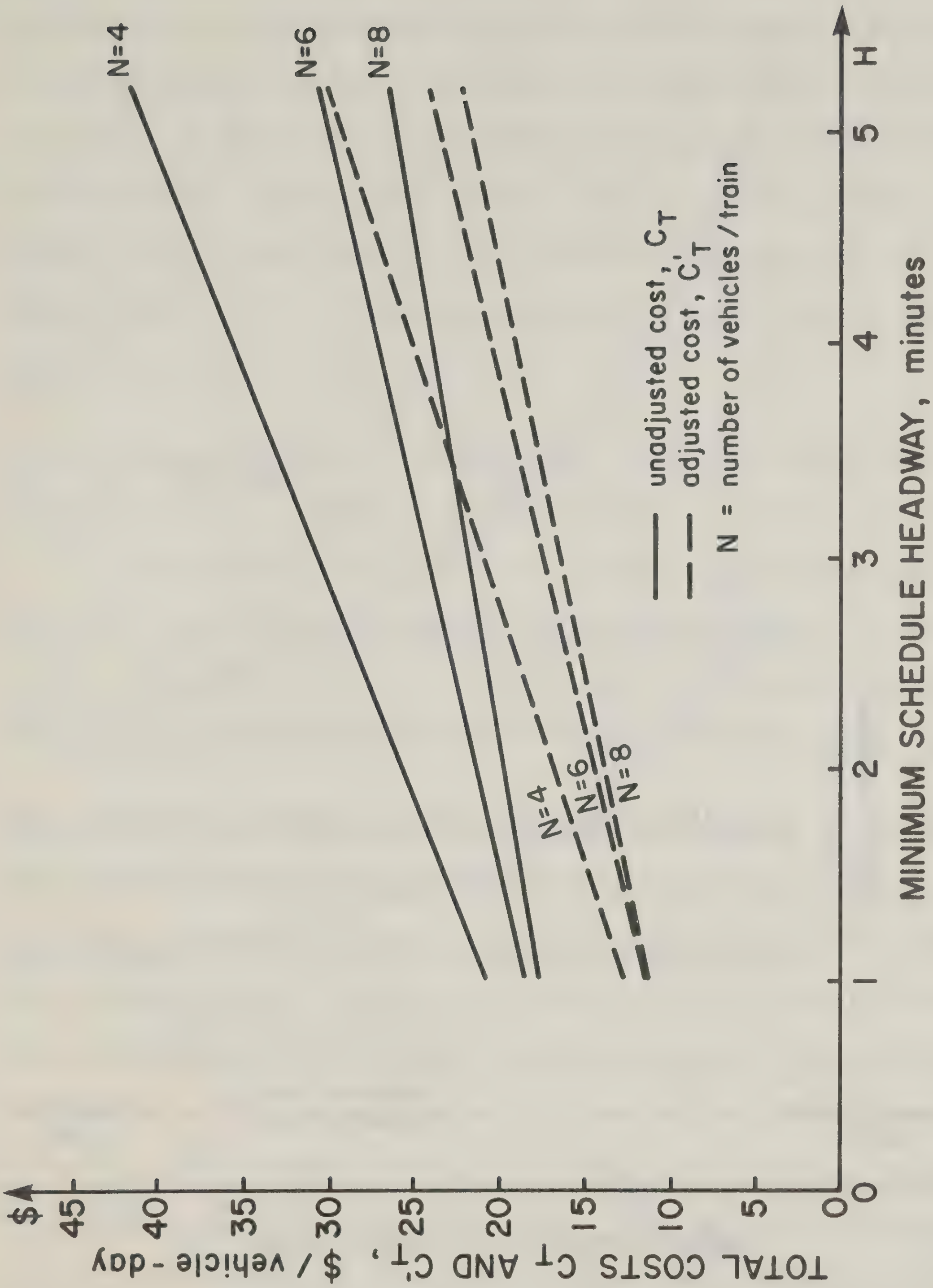


Figure 35; Total ATO-Related Incremental Cost (C_T & C'_T) vs. Minimum Scheduled Headway (H)

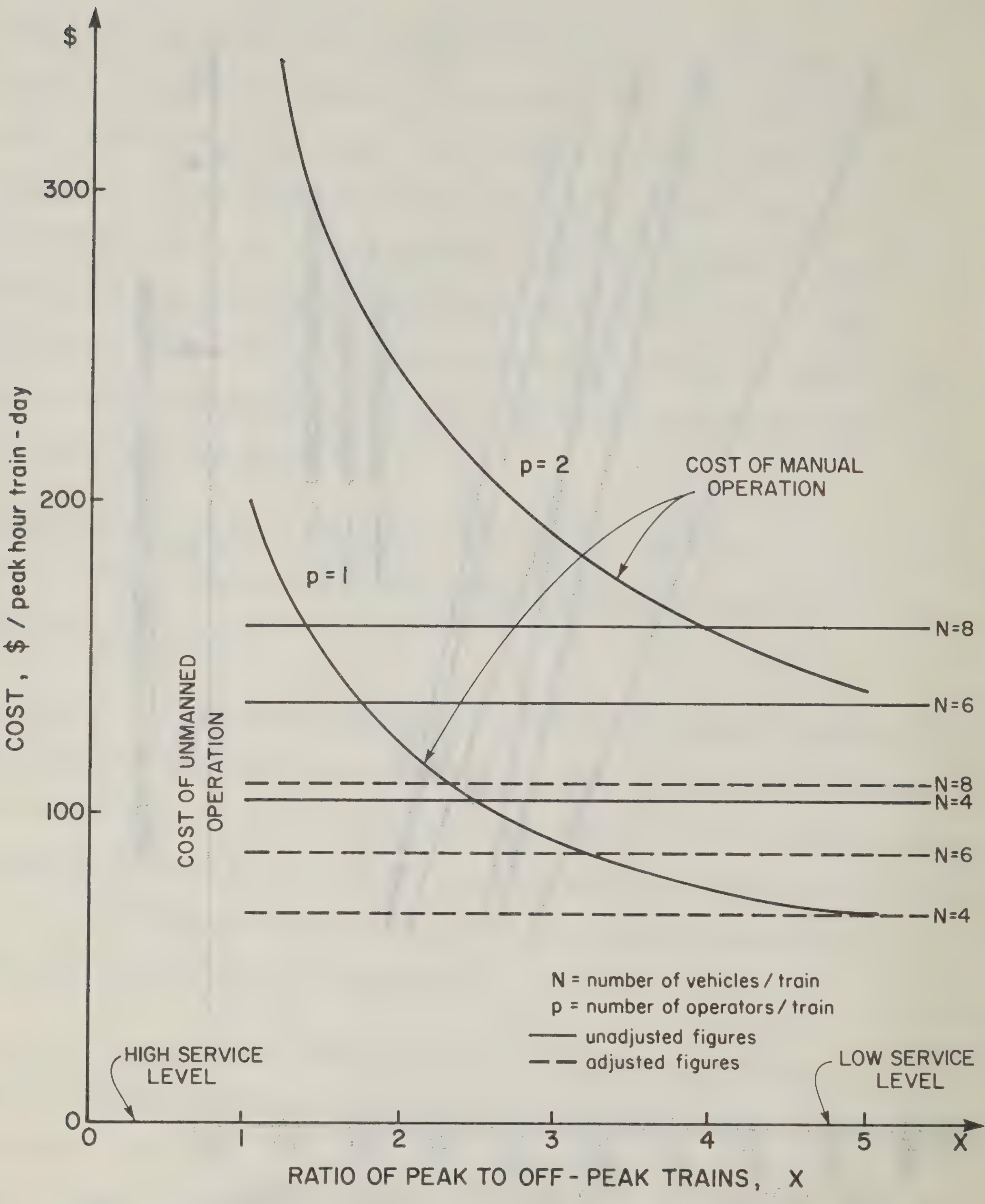


Figure 36; Cost Comparison, Manual vs. Unmanned Operation
(ATO-Related Cost Only)

headways and a 5 min off-peak headways, the ratio of peak to off-peak trains, X, is 5 : 2 or 2.5. As Figure 36 indicates, at this service level (n = 2.5), the cost of unmanned operation is higher than the cost of manual operation with a single driver for all train sizes greater than 4. If, however, the off-peak train size is adjusted to correspond to the off-peak demand level, the cost of unmanned operation is lower than the cost of manual, single-driver operation.

The cost of 2-man operation appears to be always higher than the cost of unmanned operation. As the ratio of peak trains to off-peak trains approaches 1 (i.e., equal service frequency during the whole operational day), unmanned operation shows significant savings with respect to manual operation of trains. Since increased service level, leads to increased ridership, the percentage saving is expected to be greater than that shown in Figure 36.

The previous comparisons were between manned and unmanned train operations. Figure 37 shows the incremental ATO-related costs for vehicles operating with manual back-up, i.e.,with at least one operator on-board. For this case, the ATO-related costs are lower than the cost for unmanned operations because of the operator's ability to resume operations after an ATO failure. This back-up ability reduces the service interruption costs significantly. The actual cost reduction, however, is likely to be less than predicted by this analysis when all other cost elements, such as decreased vehicle capacity due to provision for manual equipment, cost of manual equipment, etc., are all properly considered. Driver costs, of course, are in addition to the incremental ATO costs. It should be also kept in mind that the figures have been derived for 2 min minimum headway operations. For higher minimum headways all ATO costs are shifted upward. At a 5 min

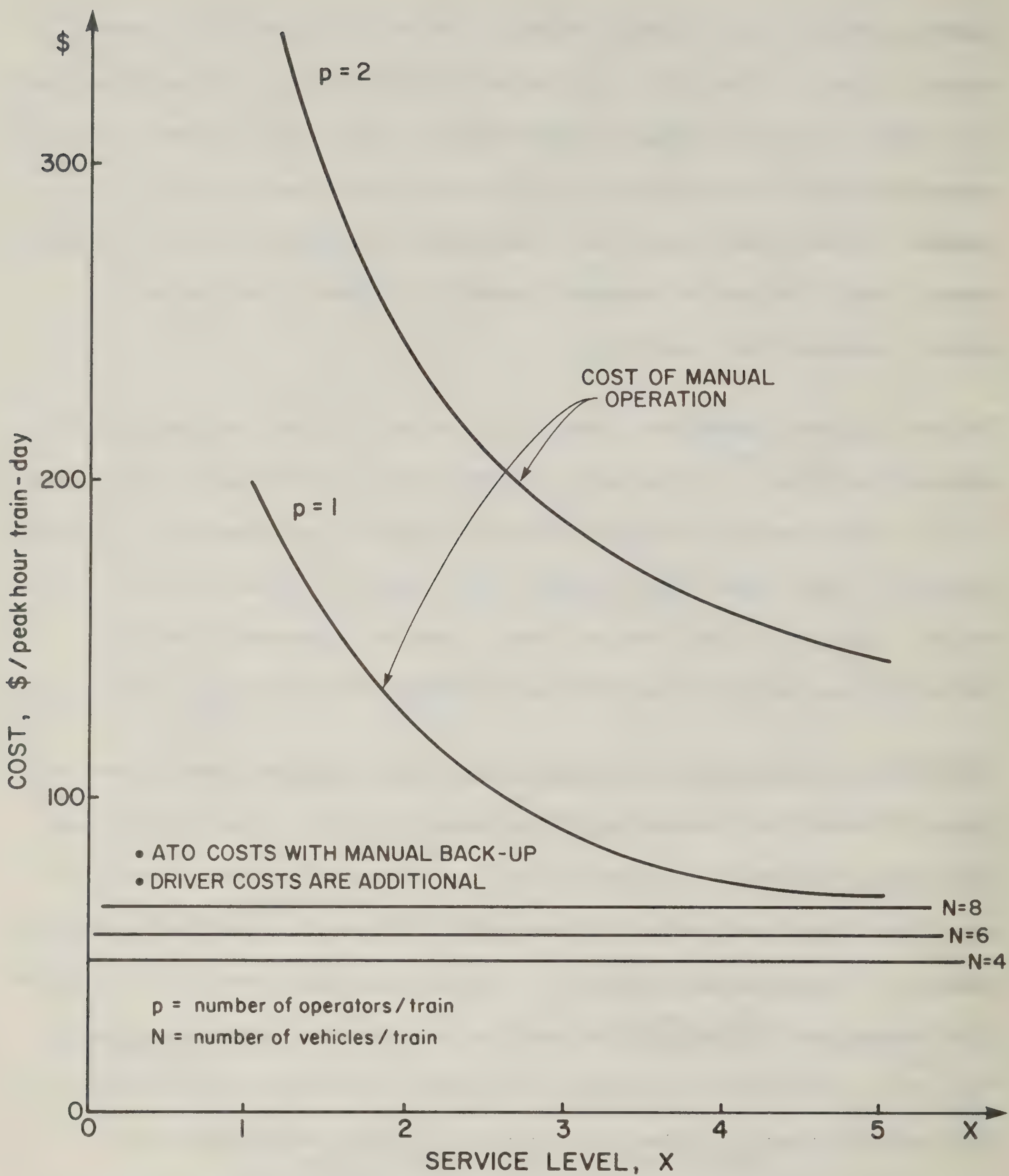


Figure 37; Cost of ATO with Manual Back-up Compared to Manual Only (Minimum Headway = 2 min.)

scheduled headway, the incremental ATO-related cost for operations with manual back-up are in the order of $100 per day for a 4-car train.

From this it may be concluded that insofar as ATO brings about a reduction of the train crew without eliminating the last crew member, it is cost-beneficial if the minimum scheduled headways are lower than 4 min, and the train size between 4 and 6 vehicles. ATO is also justified if the service frequency in the off-peak hours is to be retained at a high level.

Implicitly, it has been assumed in the analysis that the safety and security of the system has not been jeopardized by the reduction of the train crew. This assumption has not been verified by any currently-operating system, however, indications are that the system safety and security is not significantly affected by the number of on-board personnel. AIRTRANS and the MORGANTOWN system, as well as studies on the VAL and ARAMIS systems, indicate that neither real or perceived security are significantly changed under unmanned operations. This finding is strengthened by the fact that only 10 to 15% of all assaults in a transit system occur on-board the vehicle, the remaining and majority of the incidents occur outside the vehicle in the station proper - the platform, stairs, parking lots, and access routes.

The discussion was deliberately restricted to personnel cost comparisons, first of all because over 65% of transit system costs are made up of wages, salaries, and benefits and, secondly, because most other impacts of automation are difficult to quantify in a general manner. The matrix in Table 15 summarizes the cost comparison.

Table 15; Comparison Matrix

Before Introduction of ATO \ After Introduction of ATO	2-Man Crew	1-Man Crew	Unmanned
2-Man Crew	No Benefit	Beneficial	Highly Beneficial
1-Man Crew	—	No Benefit	May Be Beneficial

7.3.7 Summary

1/ In comparison to a baseline system consisting of cab signalling, overspeed protection, and a minimum communication system, the incremental cost of ATO can be justified under the following conditions:

- ATO results in the reduction of the train crew by at least one man, without eliminating the last crew member.
- ATO results in unmanned operations at short headways and high service levels with off-peak train sizes adjusted to meet the required capacity.

2/ The cost-effectiveness of ATO without human back-up in comparison with manual operation increases with increasing service frequency and increasing service level.

3/ Adjusting the off-peak train size to the capacity demand can dramatically improve the cost-effectiveness of automated train operation, provided the cost of manipulating the vehices can be kept low.

4/ The ATO-related costs appear to be less sensitive than manual operating costs to possible wage increases.

5/ Reliability appears to be a key issue in unmanned train operations. Recovery from failures in particular is a complex problem in fully automated systems which penalizes the automated operation.

8. CONCLUSIONS

1/ Constraining factors rarely permit the full realization of automatic train control potential. Full realization of automatic train control benefits can only come as a result of an appropriate and systematic approach which considers the requirements for automatic operation very early in the system design.

2/ Operational experience has proven the worth of automatic train protection (ATP) in rapid transit applications. ATP leads to safe and smooth operation. The current thinking in rapid transit train protection considers cab signalling with overspeed protection and a low performance speech/data link, as "minimum" requirement for new system applications.

3/ Even though cab signalling with overspeed protection provides a high level of safety, one of its great benefits is that it provides the basis for transition to higher levels of automatic train control. It provides experience, training, and familiarity to the operators, all of which are prerequisites for proper and beneficial application of automatic train control in rapid transit systems.

4/ Automatic Train Operation (ATO) appears to be justifiable on a cost basis if its introduction results in the elimination of at least one crew member without eliminating the last crew member, (e.g., a reduction from 2 to 1). See also Comparison Matrix (Table 15).

5/ Unmanned train operation in rapid transit systems appears to be appropriate only at short headways (approximately 1 to 2 min) and high off-peak service frequencies.

6/ The ATO-associated costs are higher for unmanned operations than for operations with a crew on-board. The cost difference results largely from the requirement to make unmanned and manned operations equally

reliable. With an operator on-board the effect of ATO failures is readily minimized, whereas in the case of unmanned service, special and costly measures are necessary.

7/ For systems which require high capacities, automatic train operation can result in 5 to 10 s headway reduction. The achievement of higher headway reductions is constrained by the turn-around at the end terminals.

8/ For systems required to operate close to capacity, the 5 to 10 s gain in headway due to ATO can be significant.

9/ If increased capacity is not a requirement, ATO does not appear to offer performance benefits in a rapid transit environment which could justify its cost.

10/ Automatic train control equipment is not less reliable than other train subsystems. The effect of train control failures, however, appears to be more pronounced.

11/ Improved reliability and better right-of-way protection for fully-automated train operations are necessary prerequisites for the successful application of automatic train operation systems.

12/ Automatic train supervision (ATS) systems are not fully developed. Insufficient operational experience exists to evaluate or judge their cost-effectiveness.

1. Theoretical Minimum On-Line Station Headway

1.1 Fixed-Block System

In a fixed-block system, the minimum separation between trains is equal to the braking distance corresponding to the maximum design operating speed (plus margin) and one block.

The minimum station headway for fixed-block systems is thus composed of five components:

1/ The station dwell time.

2/ The station clear-out time of train 1.

3/ The time required by train 2 (the following train) to travel a block length.

4/ The time required by train 2 to travel and decelerate the safe braking distance to a stop.

5/ The train control equipment reaction times.

Figure A1 depicts the situation when train 1 has just cleared the station.

Assuming that the minimum separation (one safe braking distance plus one block) is equal to n blocks and the block length is d, then

$$\text{Safe braking distance} = (n - 1) = k\,\frac{V^2}{2b} \qquad \text{(A-1)}$$

where *k = K - factor $\geqslant 1$

V = approach speed

b = braking rate

Note that the block length decreases as n increases.

---

* A ratio of train separation to braking distance; $k \geqslant 1$, brickwall stop; $k < 1$, soft stop.

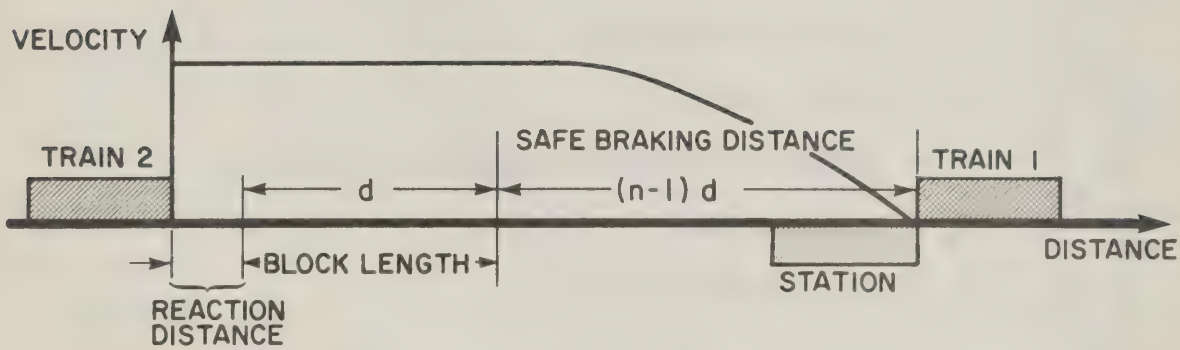


Figure A1; Station Headway

The minimum station headway H for a given approach speed can be calculated to be:

$$H = \begin{cases} \frac{V}{2a} + \frac{L}{V} + \frac{V}{2b}\left[\frac{k}{n-1} + k + 1\right] + T_r + T_d & V \leq \sqrt{2aL} \\ \sqrt{\frac{2L}{a}} + \frac{V}{2b}\left[\frac{k}{n-1} + k + 1\right] + T_r + T_d & V \geq \sqrt{2aL} \end{cases} \qquad \text{(A-2)}$$

where V = approach speed

a = nominal acceleration

b = service braking rate

L = train length

n = number of blocks of separation distance ≥ 2

T_r = reaction time

T_d = station dwell time

There is an optimum approach speed at which the minimum station headway obtains its optimum value. The optimum approach speed obtained by differentiating (A-2) is

$$H = \sqrt{\frac{2abL}{b + a\left(\frac{k}{n-1} + k + 1\right)}} \qquad \text{(A-3)}$$

The corresponding optimum mimimum station headway is

$$H_o = \sqrt{\frac{2L\left[b + a\left(\frac{k}{n-1} + k + 1\right)\right]}{ab}} + T_r + T_d \qquad \text{(A-4)}$$

The graphs of headway vs approach velocity (Figure A2 to A13) are plotted with $a = b = 0.8, 1.0, 1.2\ m/s^2$; $V = 5$ to $50\ m/s$; $L = 15, 50, 100, 150\ m$; $T_r = 1\ s$; $T_d = 15\ s$, and $k = 1.35$.

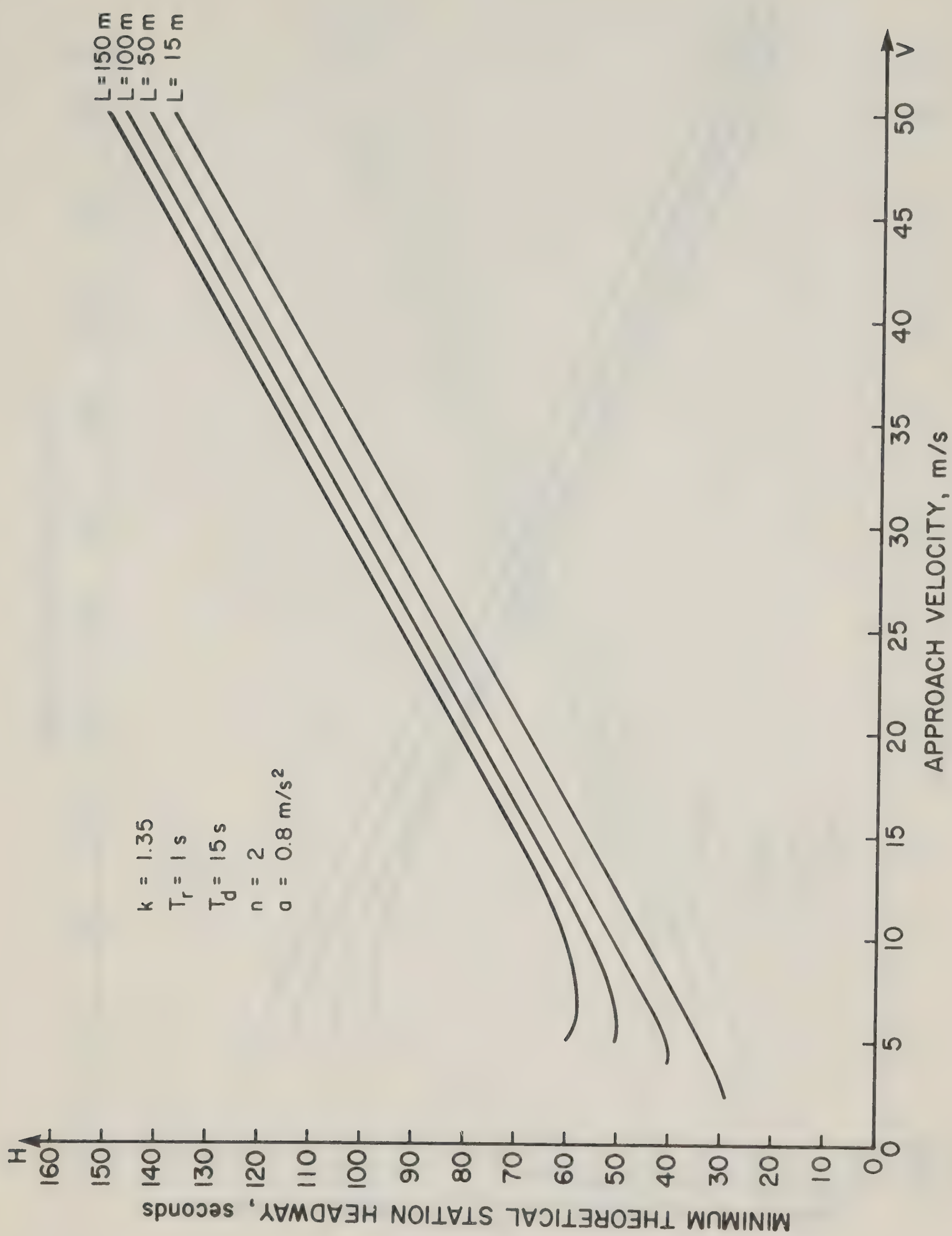


Figure A2; The Effect of Deceleration Rate and Train Length on the Relationship Between 2-Block Station Headway and Approach Velocity (Fixed-Block System)

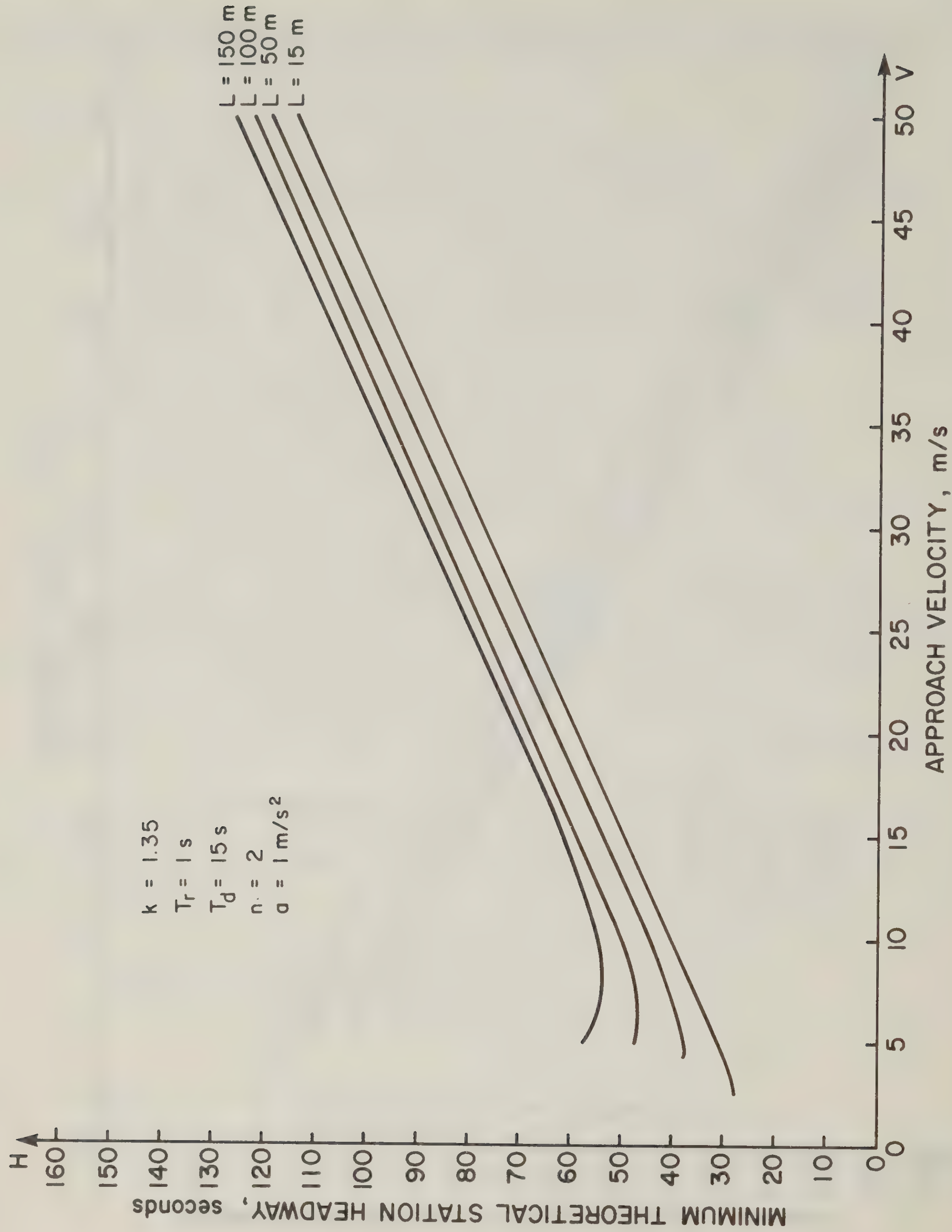


Figure A3; The Effect of Deceleration Rate and Train Length on the Relationship Between 2-Block Station Headway and Approach Velocity (Fixed-Block System)

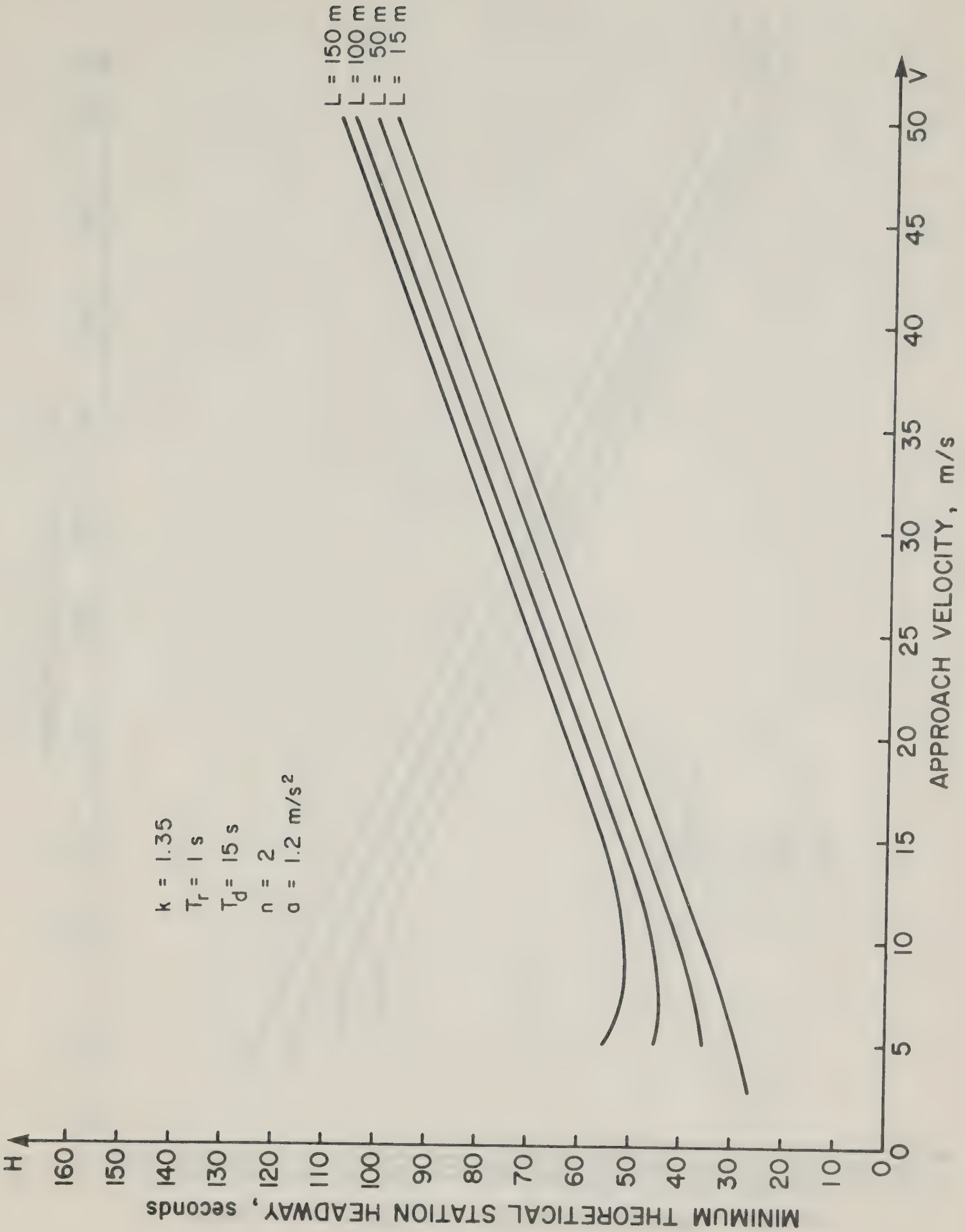


Figure A4; The Effect of Deceleration Rate and Train Length on the Relationship Between 2-Block Station Headway and Approach Velocity (Fixed-Block System)

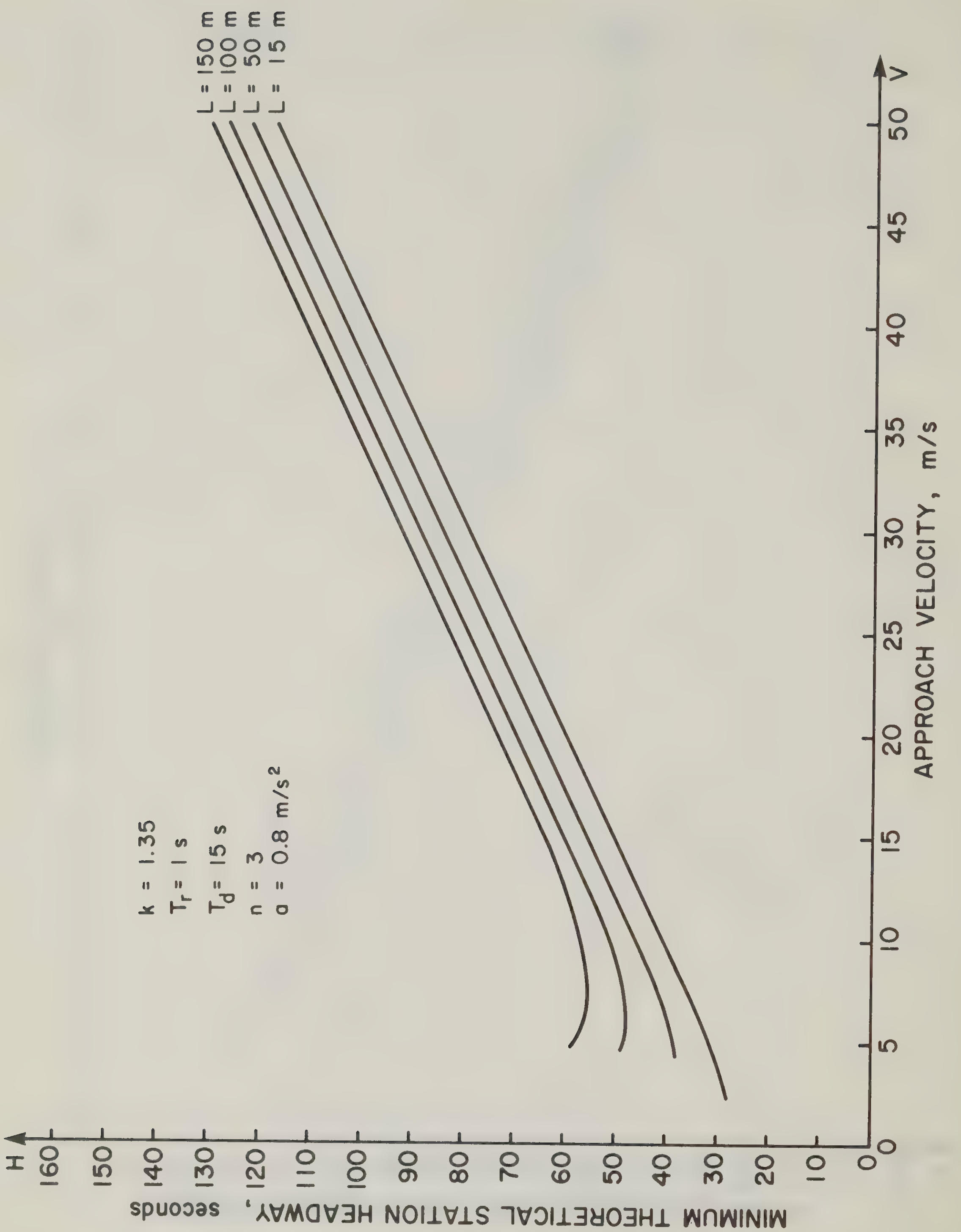


Figure A5; The Effect of Deceleration Rate and Train Length on the Relationship Between 3-Block Station Headway and Approach Velocity (Fixed-Block System)

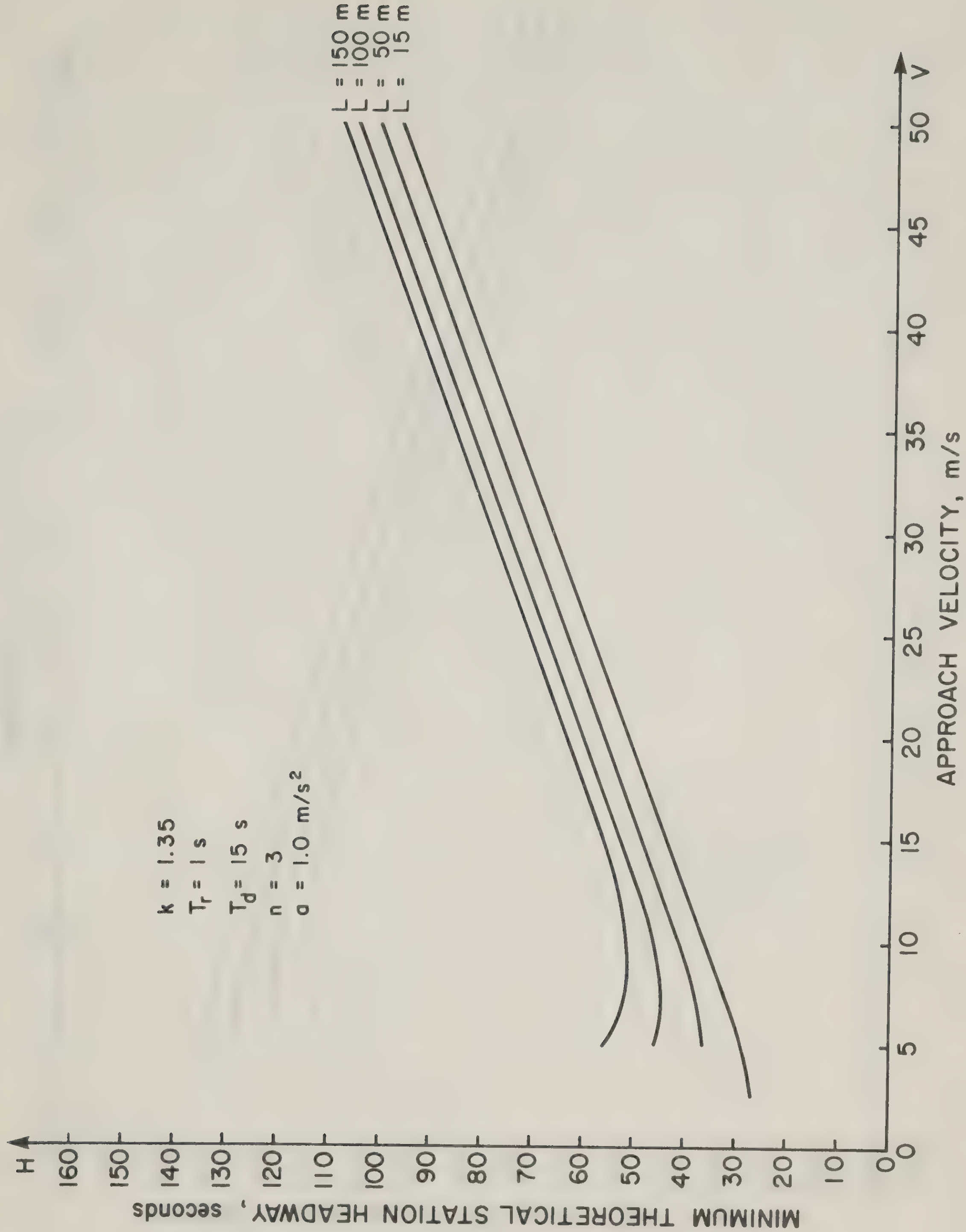


Figure A6; The Effect of Deceleration Rate and Train Length on the Relationship Between 3-Block Station Headway and Approach Velocity (Fixed-Block System)

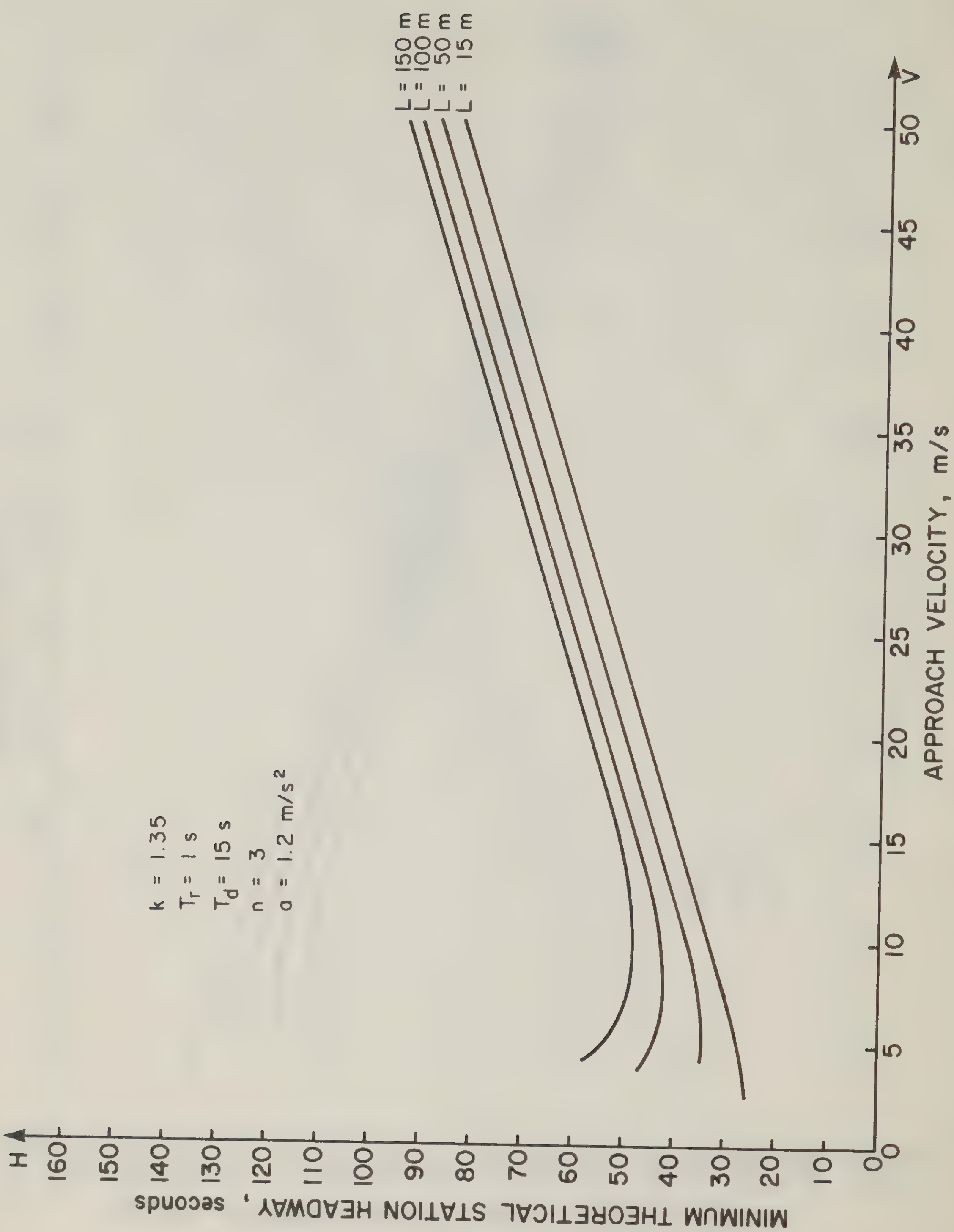


Figure A7; The Effect of Deceleration Rate and Train Length on the Relationship Between 3-Block Station Headway and Approach Velocity (Fixed-Block System)

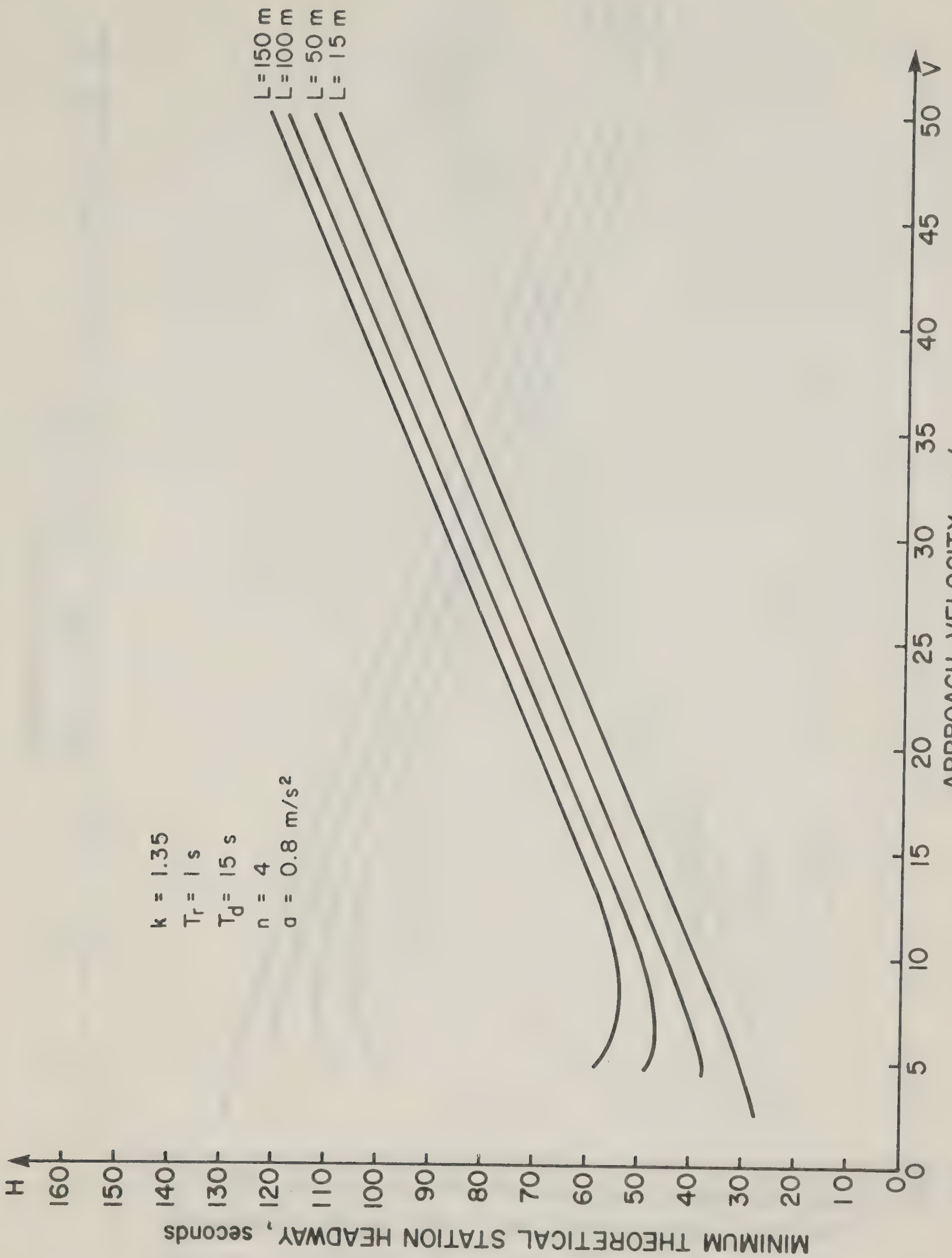


Figure A8; The Effect of Deceleration Rate and Train Length on the Relationship Between 4-Block Station Headway and Approach Velocity (Fixed-Block System)

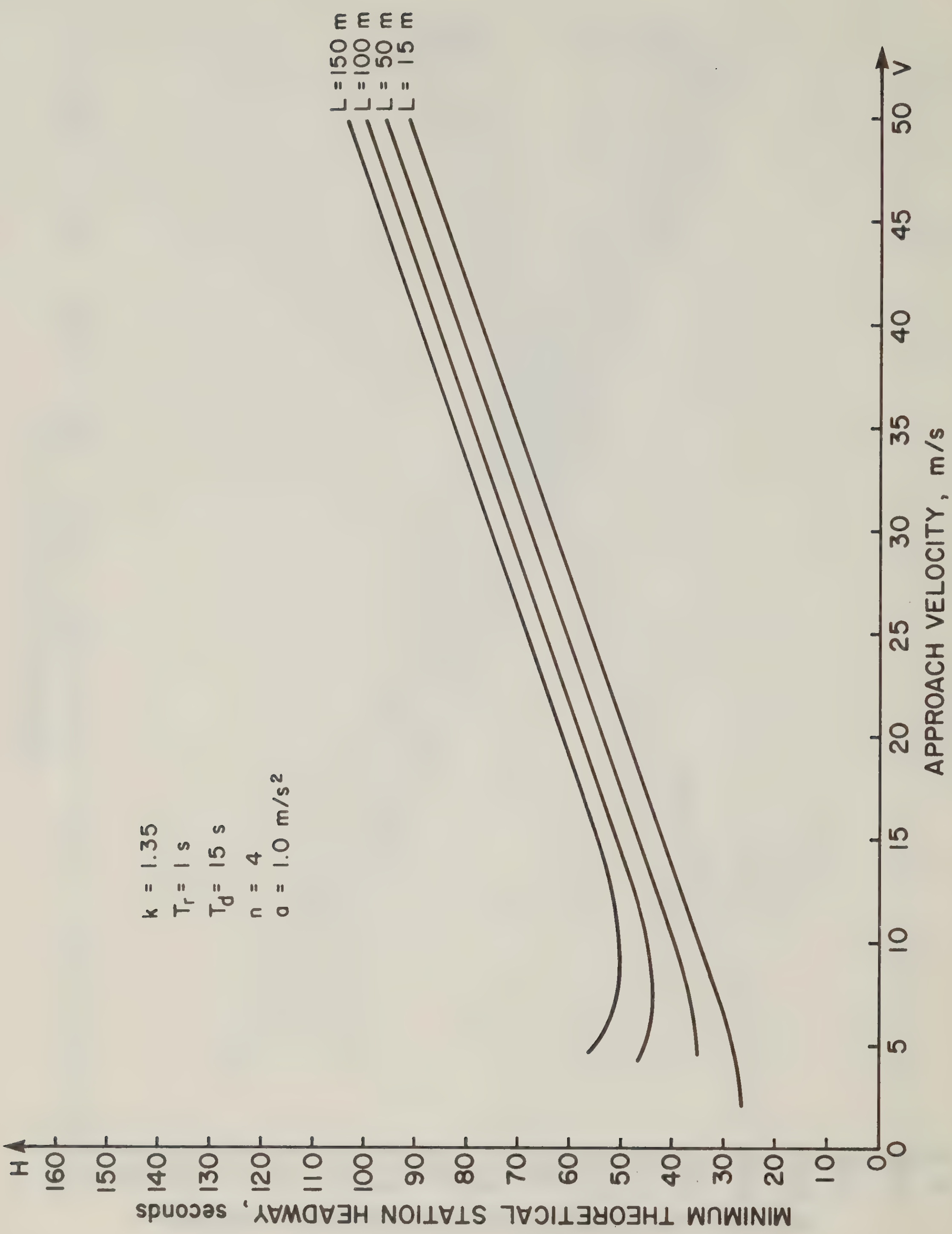


Figure A9; The Effect of Deceleration Rate and Train Length on the Relationship Between 4-Block Station Headway and Approach Velocity (Fixed-Block System)

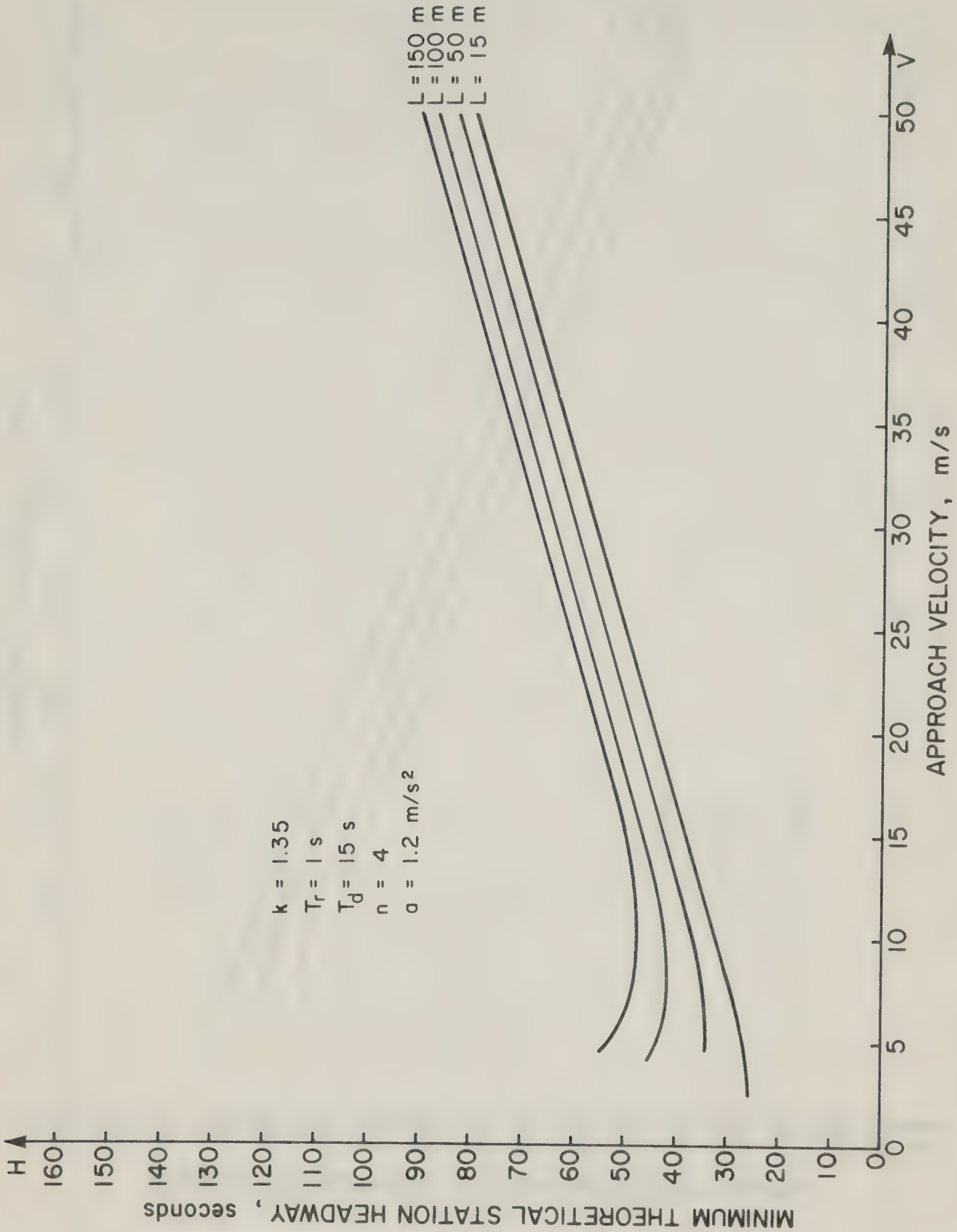


Figure A10; The Effect of Deceleration Rate and Train Length on the Relationship Between 4-Block Station Headway and Approach Velocity (Fixed-Block System)

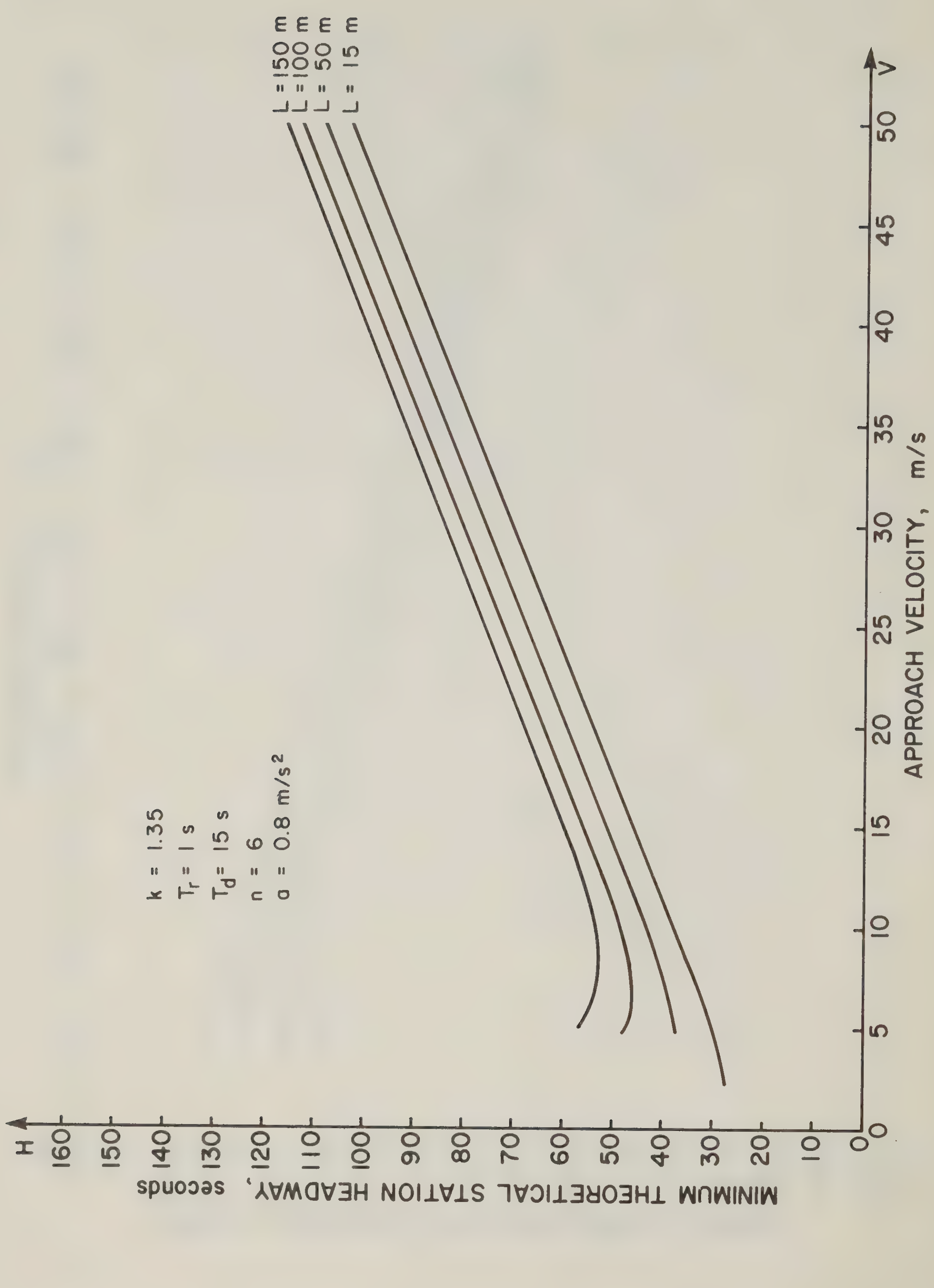


Figure A11; The Effect of Deceleration Rate and Train Length on the Relationship Between 6-Block Station Headway and Approach Velocity (Fixed-Block Systems)

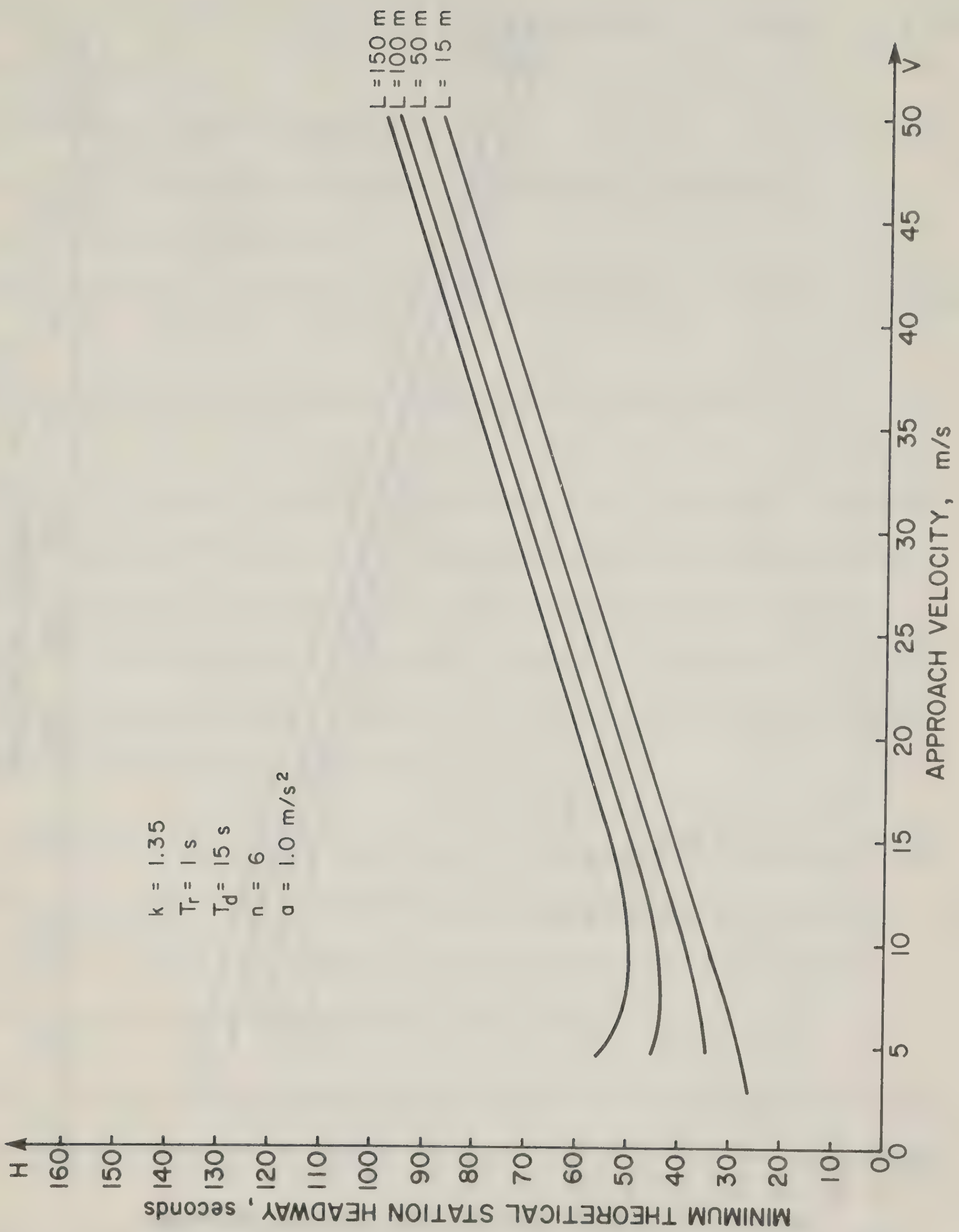


Figure A12; The Effect of Deceleration Rate and Train Length on the Relationship Between 6-Block Station Headway and Approach Velocity (Fixed-Block System)

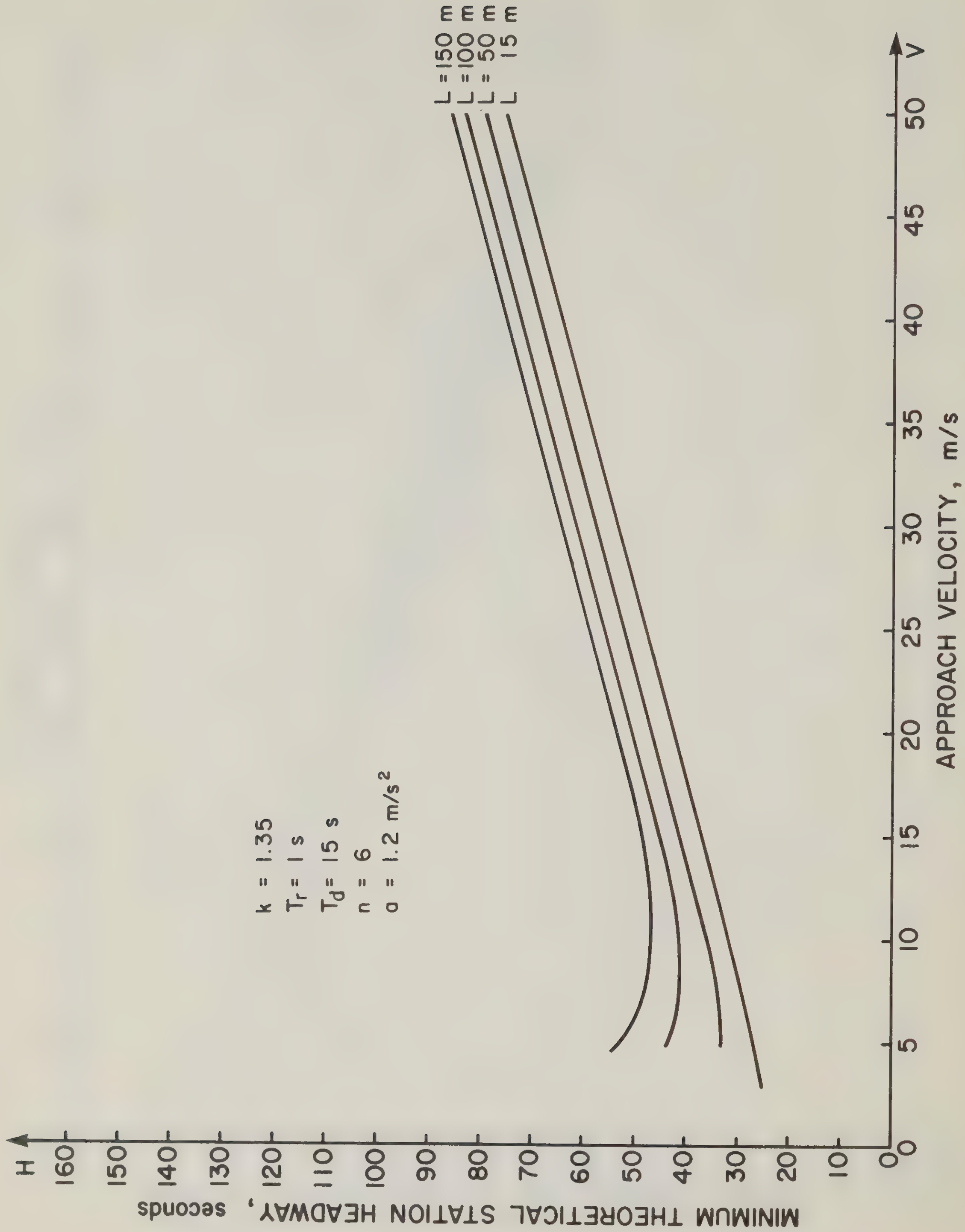


Figure A13; The Effect of Deceleration Rate and Train Length on the Relationship Between 6-Block Station Headway and Approach Velocity (Fixed-Block System)

It may be observed from Equation A-2 that headway can be decreased for a given approach velocity by

1/ Increasing n (shorter blocks).

2/ Increasing the acceleration and/or deceleration capability.

3/ Decreasing train length L.

4/ Decreasing dwell time T_d by improving the passenger loading or unloading time.

5/ Decreasing T_r and k by having automatic train operation.

Shorter blocks mean more wayside equipment and hence higher cost, since the total wayside equipment costs are inversely proportional to block length. Moreover, there is a limit to which blocks can be shortened because of engineering difficulties with very short blocks. For instance, the block boundaries are uncertain both electrically and mechanically due to signal overlap and construction tolerances.

For steel-wheel-on-steel-rail technology, the acceleration and deceleration capability is limited to about 2 m/s^2 for light rail vehicles. There is also a limit to which train length can be decreased, as the passenger carrying capacity also decreases with train length.

The above minimum station headway equation gives the lower bound for station headways for fixed-block signalling systems.

In order to make use of the automatic trip stop feature, almost all rapid transit systems employ the 2-block indication signalling plus 1-block overlap system, a discussion of which is found in Urban Rail Transit by Lang and Soberman (37). They give the theoretical minimum station headway as:

$$H = \frac{L}{V} + \frac{V}{2a} + \frac{(1 + 3k)V}{2b} + T_r + T_d \tag{A-5}$$

The corresponding optimum speed and minimum station headway are:

$$V_o = \sqrt{\frac{2abL}{b + a(1 + 3k)}} \tag{A-6}$$

$$H_o = \sqrt{\frac{2L\,[b + a(1 + 3k)]}{ab}} \tag{A-7}$$

Graphs of headway vs speed plotted with the same parameters used in plotting Equation A-2 are shown in Figures A14 to A16.

2. Moving-Block System

A description of the moving-block system can be found in Chapter 5, Section 5.2 of this report.

The theoretical minimum station headway for the moving-block system can be obtained by letting $n \to \infty$ as in Equation A-2. Therefore

$$H = \begin{cases} \dfrac{V}{2a} + \dfrac{L}{V} + \dfrac{V}{2b}(k + 1) + T_r + T_d & V \leq 2aL \\ \\ \sqrt{\dfrac{2L}{a}} + \dfrac{V}{2b}(k + 1) + T_r + T_d & V \geq 2aL \end{cases} \tag{A-8}$$

The corresponding optimum line speed and minimum station headway are

$$V_o = \sqrt{\frac{2abL}{b + a(k + 1)}} \tag{A-9}$$

and

$$H_o = \sqrt{\frac{2L\ \ b + a(k + 1)}{ab}} \tag{A-10}$$

The plots of headway vs speed are shown in Figures A17 to A19 with $k = 1$.

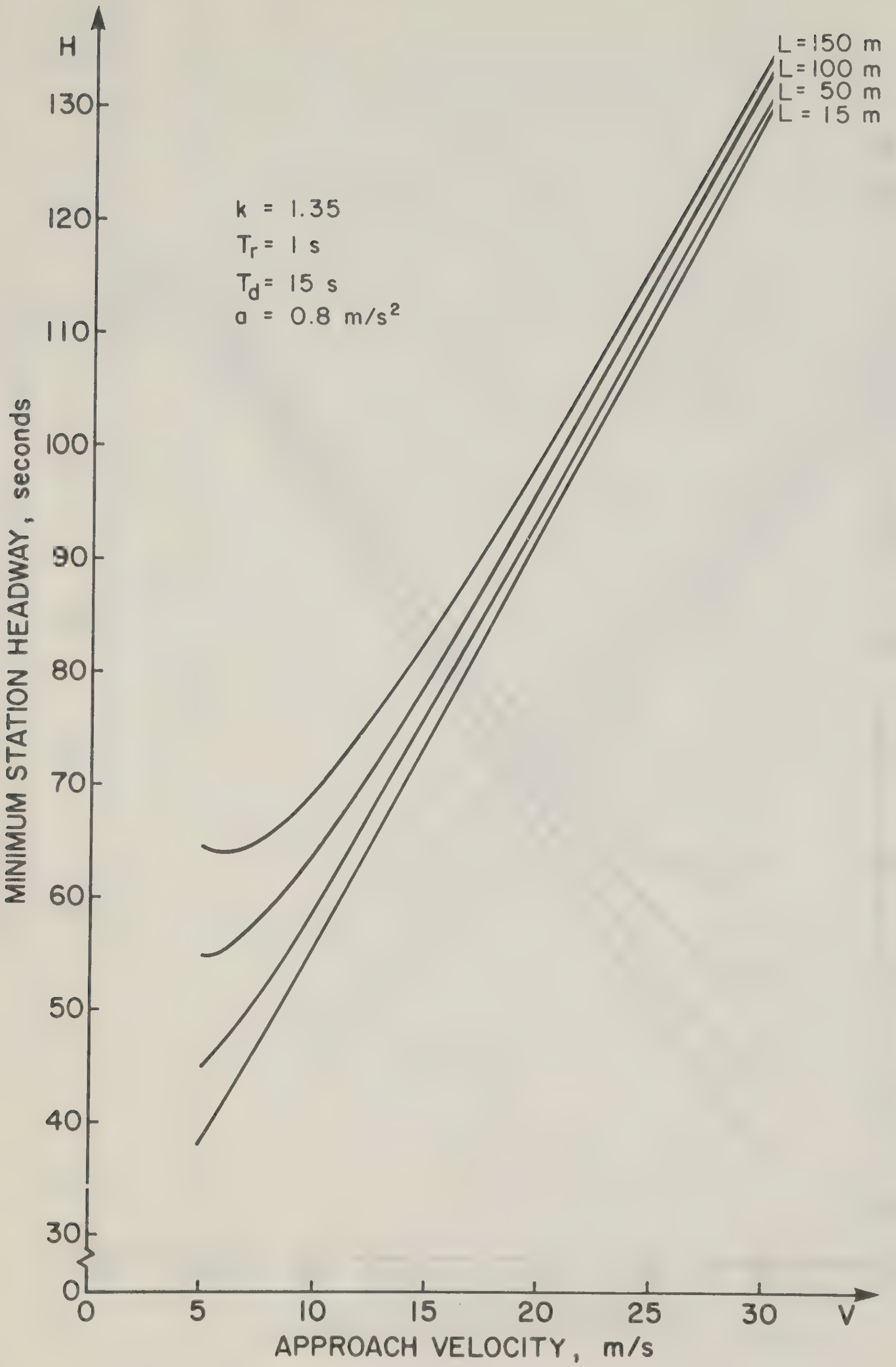


Figure A14; The Effect of Deceleration Rate and Train Length on the Relationship Between Station Headway and Approach Velocity (Two-Block Signalling Plus One-Block Overlap System)

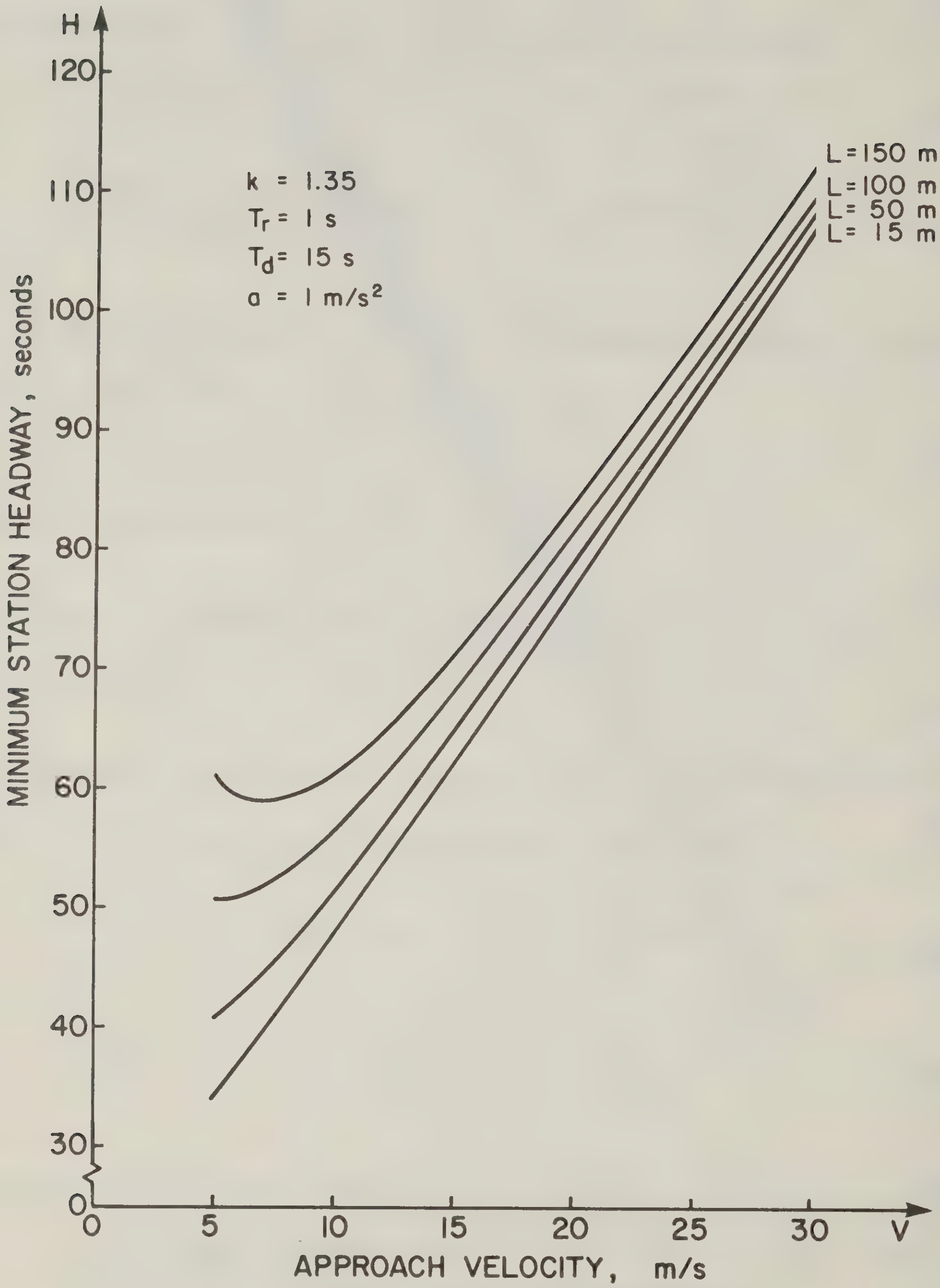


Figure A15; The Effect of Deceleration Rate and Train Length on the Relationship Between Station Headwayway and Approach Velocity (Two-Block Signalling Plus One-Block Overlap System)

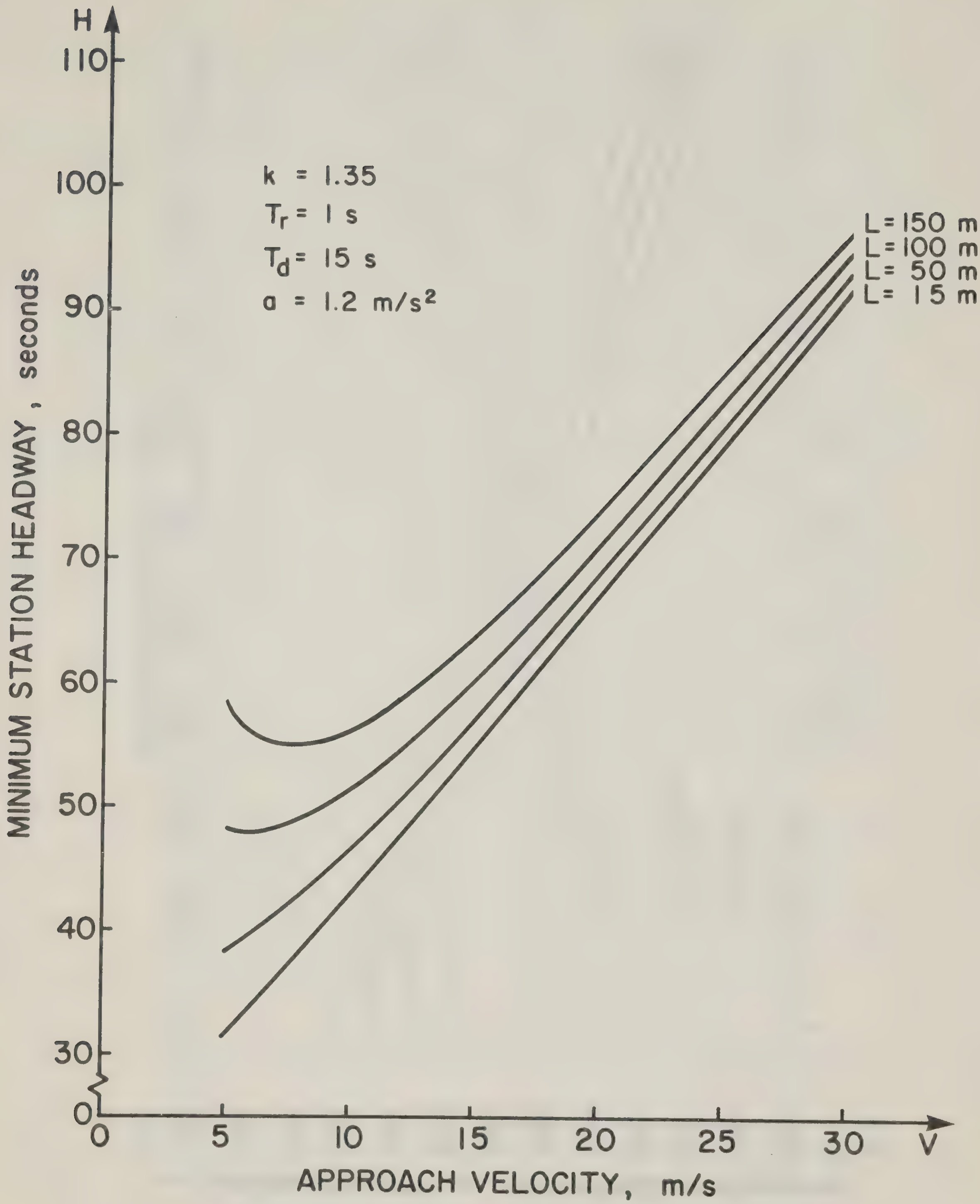


Figure A16; The Effect of Deceleration Rate and Train Length on the Relationship Between Station Headway and Approach Velocity (Two-Block Signalling Plus One-Block Overlap System)

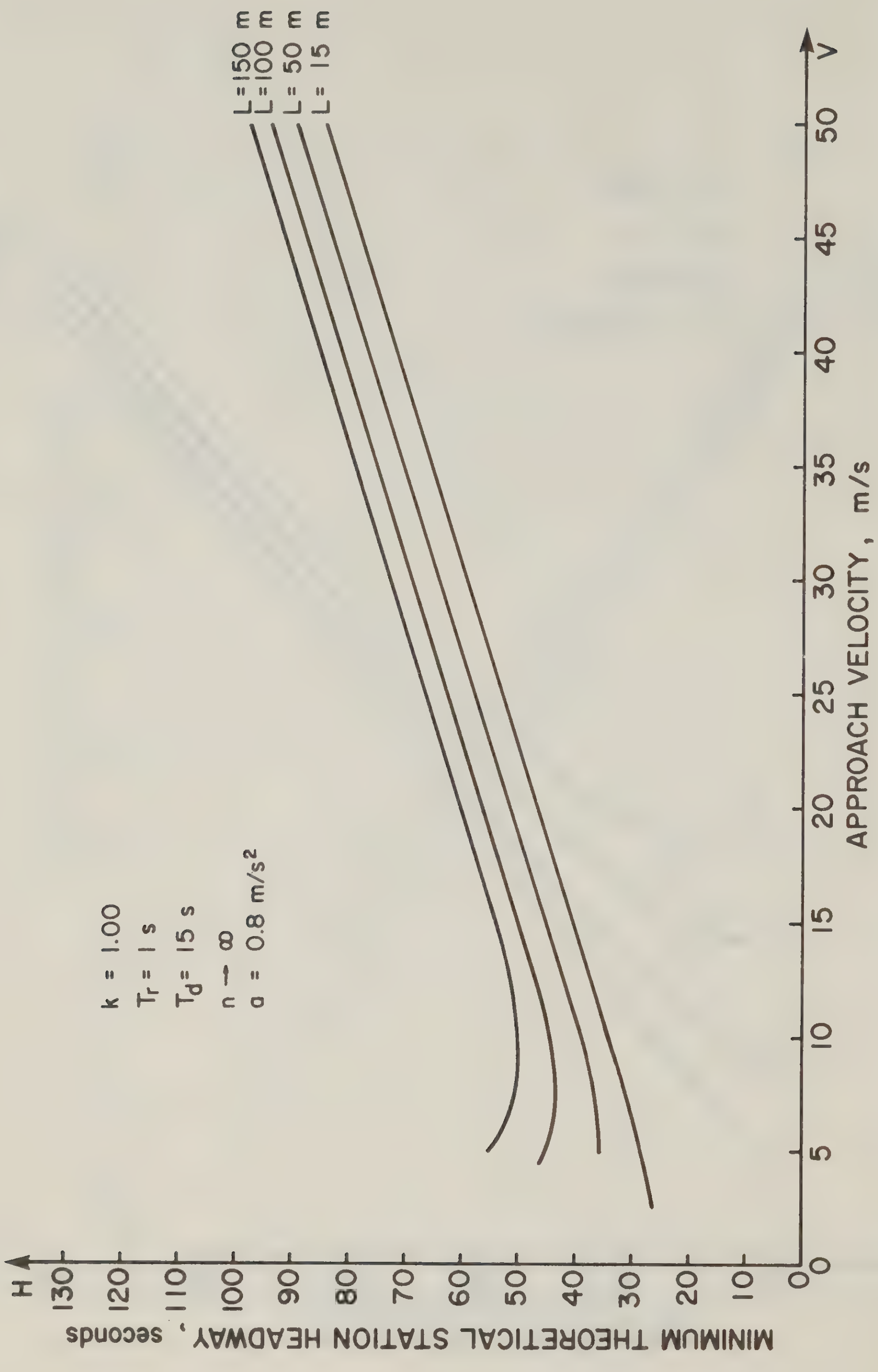


Figure A17; The Effect of Deceleration Rate and Train Length on the Relationship Between Station Headway and Approach Velocity (Moving Block System)

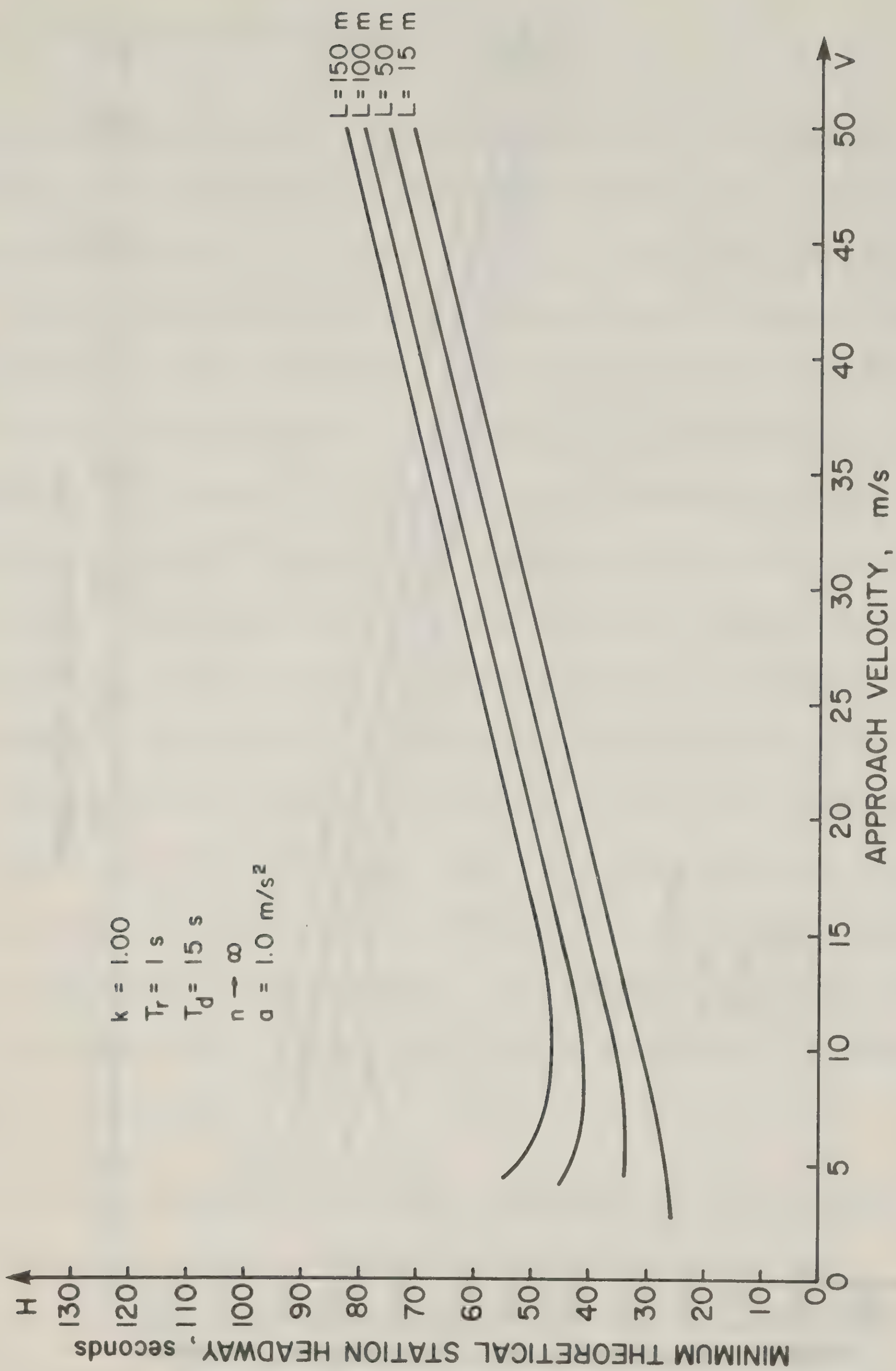


Figure A18; The Effect of Deceleration Rate and Train Length on the Relationship Between Station Headway and Approach Velocity (Moving Block System)

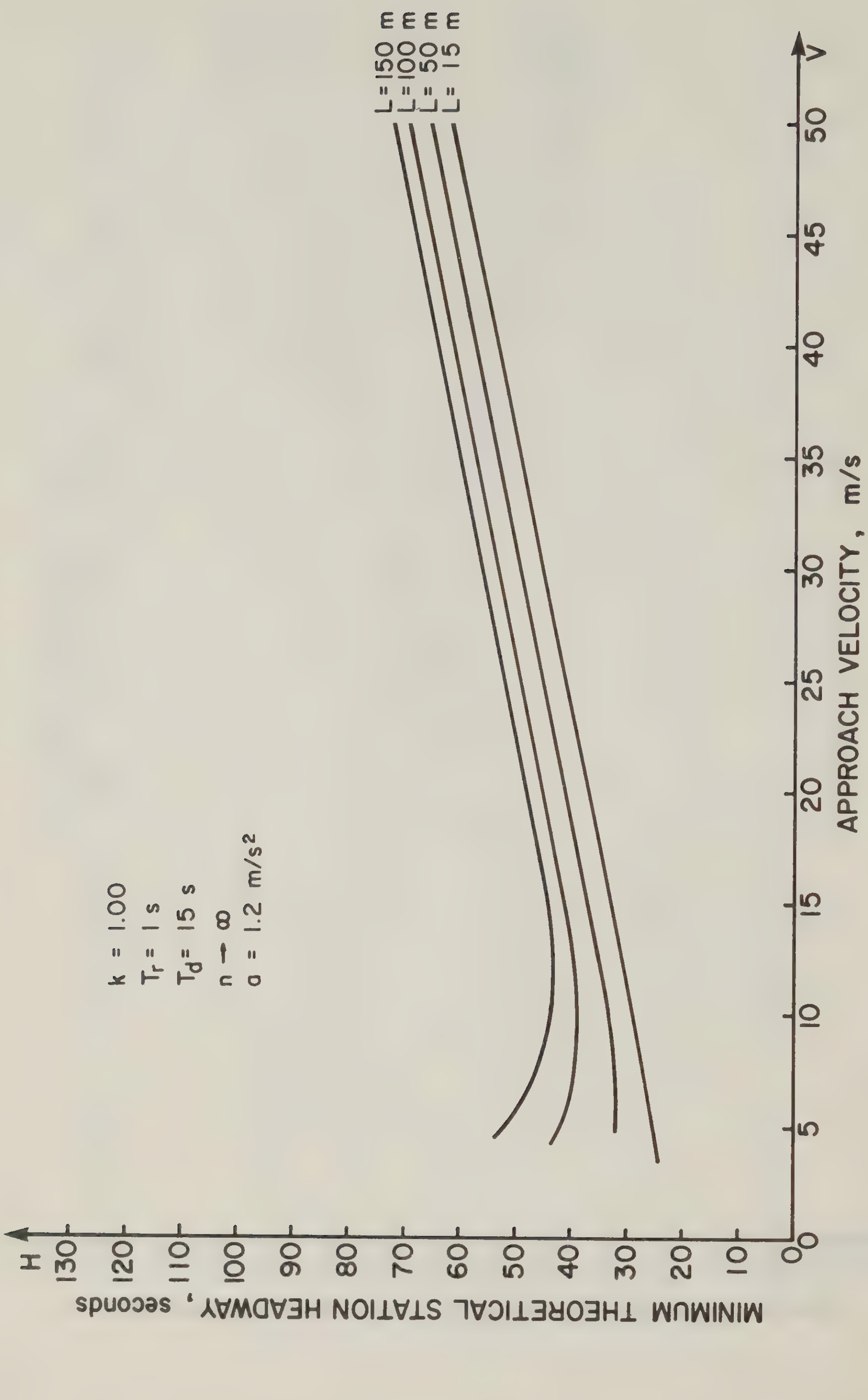


Figure A19; The Effect of Deceleration Rate and Train Length on the Relationship Between Station Headway and Approach Velocity (Moving Block System)

3. Practical Headway Considerations

The above headway equations have been derived from fully deterministic procedures based on the calculation of the minimum time interval between two successive trains arriving at the same station. In real-life operation, however, this theoretical minimum station headway cannot be achieved even if the station is the most restrictive physical bottleneck on the track. The reason is that there are stochastic variations in the parameters as well as random delays in the system. For instance, during peak-hour operation, there is substantial variation in station dwell time for loading or unloading. The presence of random variations in driver acceleration and deceleration also accounts for another stochastic variation in the system. Therefore, if the system is scheduled to operate at the theoretical minimum fixed headway, and if the first train stops at a station for longer than the nominal dwell time value, then the second train will have to slow down at a yellow aspect or stop at a red signal caused by the first train. Similarly, the subsequent trains on the line will experience the same effect. This essentially indicates an unstable situation in which serivous delay may be encountered and frequent stop-and-go driving occurs.

In order to minimize delay and have smooth operation, the usual current practice is to add an arbitrary marginal headway based on operating experience to the theoretical minimum headway to yield the desired scheduled headway. The amount of marginal headway to be added is normally line specific, one reason being that some systems may have more speed restrictions and grades than others. The TTC has a scheduled headway of 120 s during peak periods, while its theoretical minimum headway is about 90 s.

Another consideration is that in deriving the optimum approach speed, one should take into consideration the random delays in the system as well as a reasonable trip time between stations. Normally, station headway goes up with increasing approach speed. As station headway is decreased by lowering the approach velocity, the trip time is increased.

A methodological approach to the derivation of marginal headway is to model the system as a stochastic process. Mathematical analysis and simulation studies can be performed by postulating a statistical headway distribution. A discussion of the statistical approach to the headway problem is beyond the scope of this report.

APPENDIX "B"

Energy Analysis

A very simplified analysis will be used to demonstrate the potential energy/schedule-keeping gain one can achieve by trying to keep the inter-station travel time constant.

The following notation is used:

1/	Acceleration = Deceleration = constant = a			(m/s^2)
2/	Nominal schedule speed	=	V_{nom}	(m/s)
3/	Inter-station spacing	=	S	(m)
4/	Schedule deviation	=	$\pm\Delta t$	(s)
5/	Nominal schedule inter-station trip time	=	t_s	(s)
6/	Train mass	=	m	(kg)
7/	Speed required to maintain schedule in face of schedule deviation	=	V_2	(m/s)

The two simplifying assumptions used in the analysis are:

1/ The normal trip between stations consists of:

- one acceleration phase
- one coasting phase, without losses
- one deceleration phase

2/ The nominal or schedule speed is lower than the maximum achievable speed.

The time required to accelerate from zero speed to V_{nom} is simply $\frac{V_{nom}}{a}$ (s).

The time required to stop from a velocity V_{nom} is also $\frac{V_{nom}}{a}$ (s).

Consequently, the time remaining to coast is:

$$\frac{S}{V_{nom}} - \frac{V_{nom}}{a} \quad (s)$$

The scheduled trip time between two stations can therefore be expressed as:

$$t_s = \frac{V_{nom}}{a} + \frac{V_{nom}}{a} + \frac{S}{V_1} - \frac{V_{nom}}{a} = \frac{V_{nom}}{a} + \frac{S}{V_1}$$

In case of a schedule deviation, it is possible to modulate the speed in an attempt to recover. If schedule adherence is considered most important, an attempt can be made to increase the travel speed between stations so as to recover from any delays. If energy conservation is deemed important, advantage can be taken of early trains by modulating their speeds to lower than nominal values. It is, of course, possible to incorporate both considerations in a system. In the case of a Δt (s) schedule deviation, the system will try to maintain the correct (scheduled) arrival time by changing the coasting velocity from V_{nom} to V_2, that is:

$$t_s = \frac{V_{nom}}{a} + \frac{S}{V_{nom}} = \Delta t + \frac{2V_2}{a} + \frac{S}{V_2} - \frac{V_2}{a}$$

$$V_2^2 + V_2 \left(a \cdot \Delta t - V_{nom} - \frac{aS}{V_{nom}}\right) + aS = 0$$

The speed required to adhere to schedule is:

$$V_2 = 1/2 \cdot \left(\frac{aS}{V_{nom}} + V_{nom} - a\Delta t\right) \pm 1/2 \sqrt{\left(a\Delta t - V_{nom} - \frac{aS}{V_{nom}}\right)^2 - 4\, aS}$$

The energy required to accelerate to the nominal velocity is simply:

$$(ma) \cdot 1/2 \frac{V_{nom}^{2}}{a} = \frac{m \cdot V_{nom}^{2}}{2} = E_{nom}$$

The energy required to accelerate to V_2 is:

$$E_2 = \frac{m V_2^2}{2}$$

The ratio of nominal energy E_1 to the energy required E_2 to maintain the schedule in face of a schedule deviation is:

$$\frac{E_{nom}}{E_2} = \frac{V_{nom}^2}{V_2^2}$$

Table BI shows possible energy savings for early departures (negative values of Δt) and the possible increase in energy requirements for attempting to maintain schedule in face of a delay.

Table B1; Possible Energy Savings Due to ATO

For:

$V_{nom} = 20$ m/s

$a = 1$ m/s^2

$s = 1000$ m

Δt (s)	$E_2/E_1 \times 100$	V_2 (m/s)
2	126.6%	22.5
4	139 %	23.6
6	189 %	27.5
-2	88 %	18.8
-4	79 %	17.8
-6	72 %	17

For:

$V_{nom} = 20$ m/s

$a = 1$ m/s^2

$s = 2000$ m

Δt (s)	$E_2/E_1 \times 100$	V_2 (m/s)
2	105%	20.52
4	111%	21.1
6	117%	21.7
-2	95%	19.5
-4	90%	19
-6	87%	18.7

BIBLIOGRAPHY

1. Bergmann, Dietrich R.,
"Generalized Expressions for the Minimum Time Interval Between Consecutive Arrivals at an Idealized Railway Station", Transportation Research Vol. 6, 1972, pp.327-341.

2. "Electronic Railway Interlocking",
The Engineer, Dec. 16, 1960.

3. "Cab Signalling and Speed Control for Mainline Railways",
GRS Bulletin 208, Jan. 1975.

4. "Continuous Automatic Train Control on Munich S-Bahn",
Railway Gazette, Jan. 1974.

5. McAulay, Alastair D.,
"Progress in Signalling for Track Guided Systems", Transportation Engineering Journal, Nov. 1975, pp.621-636.

6. Cunliffe, J.P.,
"A Survey of Railway Signalling and Control", Proceedings IEEE, Vol.56, No. 4, April 1968.

7. "Automatic Train Stop Device, 'S' Type Cab Warning Device",
The Nippon Signal Co. Ltd.

8. Itoh, Kikuji,
"New ATC System by Non-Insulated Track Circuits", Japanese Railways Quarterly Report, Vol. 13, No. 3, 1972.

9. "Automatic Train-Speed Control Equipment",
ACEC Review, No. 2, 1970.

10. Kawanabe, Hajime,
"Automatic Train Control (ATC) - Present and Future", Japanese Railways Quarterly Report, Vol.10, No. 1, 1969.

11. Prof. Dr.-Ing. Hans Fricke and Dr.-Ing. Peter Form, Brunswick U.,
"Application of Communication Techniques to a Future System of Train and Line Protection", IRCA Bulletin: Cybernetics and Electronics on the Railways, April 1966.

12. Carey, A.J.,
"Basic Track Circuit Limitations", Wabco Bulletin 261.

13. "Automatic Train Control in Rapid Transit",
U.S. Congress, Office of Technology Assessment, May 1976.

14. "Guidelines and Principles for Design of Rapid Transit Facilities",
IRT, May 1976.

15. Hadaway, H.W.,
"Fail Safe", London Transport.

16. "Safety and Automatic Train Control for Rapid Transit Systems", NTIS PB-235 492, June 1974.

17. Tappert, A., Mies, A., and Lindner, J., "Automating The Hamburg Underground".

18. Maxwell, W.W., "Disappearing Staff Force - The Pace of Automation", Railway Gazette International, Jan. 1974, pp. 9-14.

19. Vuchic, V.R., "Rapid Transit Automation and the Last Crew Member", Railway Gazette, Oct. 1973.

20. Masiewicz, A., "An Evaluation of the Technical Criteria of Train Operations for ATC Equipment", AFCET, Traffic Control and Transportation Systems, 1974, pp. 441-451.

21. Bauchman, G.W., "Step By Step Approach to Crewless Operation", Financial Analysis Journal May/June 1962, pp. 67-78.

22. Maxwell, W.W. and Ware, D.K., "Automatic Train Operation on London Transport Railways", The Institute of Locomotive Engineers Journal, Vol. 56, No. 6, 1966/67.

23. "Towards the Automatic Railway", Railway Gazette International, Aug. 1972.

24. "Washington Opts For Stand-Alone Automation", Railway Gazette International, Aug. 1972.

25. Hahn, H.J., "Automatic Operation of Urban Transit and Underground Railways", Brown Boveri Review 12, Vol. 58, Dec. 1971.

26. "Germany Automates its Rails", IEEE Spectrum, July 1974.

27. George, F., "Report on Automatic Train Operation", International Railway Congress Association Bulletin, May 1961.

28. Freehafer, J.E., "Design Philosophies in Automatic Train Control", Railway Gazette International, October 1975.

29. Quintin, W.P., "Automatic Train Control for BART", IEEE Railroad Conference, 1966.

30. McRillop, A.N., "Automatic Operation of Rapid Transit Trains" DeLeuw, Chadwick Consultants.

31. "Automatic Train Operation",
The Engineer, April 10, 1964.

32. Shirai, Yoshimi and Ishihara, Yonehiko,
"Teito Rapid Transit Authority Automatic Train Operation", Proceedings IEEE, Vol.56, No. 4, April 1968.

33. "The Analysis of Skills in Driving",
Australian Road Research, March 1964.

34. "Control Considerations for Short Headway ACGV Systems",
APL, Johns Hopkins U., October 1971.

35. Freehafer, J.E.,
"The Architecture of Command and Control Systems for Modern Rapid Transit".

36. "American Railway Signalling Principles and Practices",
AAR Communications and Signal Section, Washington, D.C., April 1955.

37. Lang, A.S. and Soberman, R.M.,
"Urban Rail Transit", MIT Press, 1964.